1972
16.2.67

152

# ACQUISITION OF SKILL

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
LIBRARY
CALCUTTA

# *Conference Participants*

Jack A. Adams, University of Illinois
Robert E. Adamson, U. S. Air Force
Earl A. Alluisi, University of Louisville
Burton G. Andreas, University of Rochester
Fred Attneave, University of Oregon
Harry P. Bahrick, Ohio Wesleyan University
William F. Battig, University of Maryland
Edward A. Bilodeau, Tulane University
Ina McD. Bilodeau, Tulane University
David Birch, University of Michigan
L. E. Bourne, Jr., University of Colorado
George E. Briggs, Ohio State University
Lee S. Caldwell, U. S. Army
R. Ernest Clark, U. S. Army
Charles N. Cofer, Pennsylvania State University
Edwin A. Fleishman, American Institutes for Research
Paul W. Fox, University of Minnesota
Morton P. Friedman, University of California
G. Robert Grice, University of Illinois
Glenn R. Hawkes, U. S. Army
Marvin J. Herbert, U. S. Army
Arthur L. Irion, Tulane University
Earl I. Jones, U. S. Navy
Marshall B. Jones, University of Florida
C. Michael Levy, University of Florida
Clyde E. Noble, University of Georgia
E. C. Poulton, Applied Psychology Research Unit, England
Karl W. Smith, University of Wisconsin
Edward R. Stearns, U. S. Army
Benton J. Underwood, Northwestern University

# ACQUISITION OF SKILL

*Edited by*

EDWARD A. BILODEAU

DEPARTMENT OF PSYCHOLOGY
TULANE UNIVERSITY
NEW ORLEANS, LOUISIANA

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

BUREAU OF EDUCATIONAL AND PSYCHOLOGICAL RESEARCH
LIBRARY
CALCUTTA

1966

ACADEMIC PRESS □ New York and London

16.2.67
1972

COPYRIGHT © 1966 BY ACADEMIC PRESS INC.
ALL RIGHTS RESERVED
NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM,
BY PHOTOSTAT, MICROFILM, OR ANY OTHER MEANS, WITHOUT
WRITTEN PERMISSION FROM THE PUBLISHERS.

151
BIL

ACADEMIC PRESS INC.
111 Fifth Avenue, New York, New York 10003

*United Kingdom Edition published by*
ACADEMIC PRESS INC. (LONDON) LTD.
Berkeley Square House, London W.1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 65-26036

PRINTED IN THE UNITED STATES OF AMERICA

S.C.E.R.T., West Bengal
Date 16.2.67
Acc. No. 1972

# List of Contributors

*Numbers in parentheses indicate the pages on which the authors' contributions begin.*

JACK A. ADAMS, *Department of Psychology, University of Illinois, Urbana, Illinois* (169)

EARL A. ALLUISI, *Department of Psychology, University of Louisville, Louisville, Kentucky* (201)

FRED ATTNEAVE, *Department of Psychology, University of Oregon, Eugene, Oregon* (483)

HARRY P. BAHRICK, *Ohio Wesleyan University, Delaware, Ohio* (351)

WILLIAM F. BATTIG, *Department of Psychology, University of Maryland, College Park, Maryland* (215)

EDWARD A. BILODEAU, *Department of Psychology, Tulane University, New Orleans, Louisiana* (315)

INA MCD. BILODEAU, *Department of Psychology, Tulane University, New Orleans, Louisiana* (255)

DAVID BIRCH, *Department of Psychology, University of Michigan, Ann Arbor, Michigan* (99)

L. E. BOURNE, JR., *Department of Psychology, University of Colorado, Boulder, Colorado* (297)

GEORGE E. BRIGGS, *Office of Research, Ohio State University, Columbus, Ohio* (411)

EDWIN A. FLEISHMAN, *American Institutes for Research, Washington, D.C.* (147)

PAUL W. FOX, *Department of Psychology, University of Minnesota, Minneapolis, Minnesota* (245)

ARTHUR L. IRION, *Department of Psychology, Tulane University, New Orleans, Louisiana* (1)

MARSHALL B. JONES, *Department of Psychiatry, J. Hillis Miller Health Center, University of Florida, Gainesville, Florida* (109)

CLYDE E. NOBLE, *Department of Psychology, University of Georgia, Athens, Georgia* (47)

E. C. POULTON, *Medical Research Council, Applied Psychology Research Unit, Cambridge, England* (361)

KARL U. SMITH, *Behavioral Cybernetics Laboratory, University of Wisconsin, Madison, Wisconsin* (425)

BENTON J. UNDERWOOD, *Department of Psychology, Northwestern University, Evanston, Illinois* (489)

# Preface

The field of motor-skill acquisition is a part of experimental psychology and one of the siblings of the learning family. Those who work in the area think of themselves as contributing more to *learning* and less to *motor* and few investigators care much about motor-skill *qua* motor-skill. Rather, the research emphasis is upon the variables and explanations of learning that will be seen in the chapters which follow.

This book contains the papers presented at a conference on "Acquisition of Skill" held March 8–12, 1965 in New Orleans, Louisiana. The conference was supported by the U. S. Army Medical Research and Development Command, Office of the Surgeon General, through Grant No. 65-G140 to Tulane University. As in the Cofer conferences on verbal behavior which preceded it, Principal authors were asked to submit chapters before conference time and Discussant authors were asked to submit reviews by conference time. The papers by the Discussants were read at the conference; those of the Principals were not. Both sets are published here with but minor changes. The greater part of the conference was spent in general discussion of the work of Principal and Discussant; no summary of this activity was recorded.

A book promised to be very useful since there was no previous one, although a number of highly regarded investigators had contributed to the area regularly. These investigators, while individually prolific, had not brought attention to the field by integrating their diverse contributions or writing general summaries. Although this helped to retard recognition, acknowledgment of motor-skills contributions to human behavior theory has nonetheless come, if slowly and recently, as, for example, in the decision to cover motor-skills acquisition in the *Annual Review of Psychology* of 1961 and in its inclusion as a category in Melton's *Categories of Human Learning* in 1964. In 1961, the writer determined that skill would have book-length coverage; the present volume is the result.

The book begins with a history of research on the acquisition of skill—the only such historical survey available. The second chapter, on selective learning, is the first substantive chapter. It is first because motor involvement is deliberately held to a minimum and the process of response selection can be more readily seen. Chapter 3 covers the topic of individual differences, the methodology of test and measurement that has been brought to bear on problems usually reserved to experimental psychology. A succession of chapters (4, 5, 6, 7) follows in which the motor literature, and some verbal as well, is covered, but in which the authors make more of a point of speaking to the issues of learning. Chapters 8 and 9 are more frankly motor, dealing with tracking and moving, respectively. Finally, the last chapter acknowledges the need to address ourselves to motor and verbal processes simultaneously and includes a critical review of the previous chapters.

It is not possible to acknowledge all the advice and counsel received. The sponsor deserves not only recognition for its financial support, but also for the painless way that Dr. Glenn R. Hawkes, Chief of the Basic Sciences Branch of the Office of the Surgeon General represented the agency. Dr. Ina McD. Bilodeau and Dr. Arthur L. Irion gave encouragement and direction during the years the conference and book were under preparation. All of the conferees provided counsel at one time or other, their cooperation producing my enduring gratitude.

*December, 1965* EDWARD A. BILODEAU

# Contents

CONFERENCE PARTICIPANTS .......... ii

LIST OF CONTRIBUTORS .......... v

PREFACE .......... vii

## *Chapter 1.* A Brief History of Research on the Acquisition of Skill 1

ARTHUR L. IRION

Introduction .......... 1
The Early Period .......... 9
The Second Period .......... 12
The Third Period .......... 19
References .......... 37

## *Chapter 2.* Selective Learning 47

CLYDE E. NOBLE

Multiple-Choice Situations .......... 47
Theory of Selective Learning .......... 56
Laws of Selective Learning .......... 73
Summary .......... 89
References .......... 90

## Comments on Professor Noble's Paper 99

DAVID BIRCH

Theory of Selective Learning .......... 99
An Expanded Domain for Selective Learning .......... 101
Conclusions .......... 106
References .......... 107

*Chapter 3.* **Individual Differences** 109

MARSHALL B. JONES

Acquisition ........ 112
Relearning and Recall ........ 124
Inheritance ........ 128
Task Classification ........ 131
Selection ........ 136
Training ........ 139
Tight Little Islands ........ 143
References ........ 144

Comments on Professor Jones' Paper 147

EDWIN A. FLEISHMAN

The Ability-Skill Distinction ........ 147
Abilities and Skill Learning ........ 148
Taxonomy of Perceptual-Motor Abilities ........ 150
Skill Acquisition ........ 158
Individual Differences in Retention ........ 163
Individual Differences and Other Learning Phenomena ........ 164
Abilities and Task Requirements ........ 164
Concluding Statement ........ 165
References ........ 165

*Chapter 4.* **Some Mechanisms of Motor Responding: An Examination of Attention** 169

JACK A. ADAMS

The Observing Response ........ 175
Anticipation ........ 182
Limitations in Human Data Processing ........ 190
Conclusions ........ 196
References ........ 197

Comments on Professor Adams' Paper 201

EARL A. ALLUISI

Attention and Research on Vigilance ........ 202
Vigilance and Arousal Theory ........ 209

Mechanics of Mechanisms of Response 210
References 211

## *Chapter 5.* Facilitation and Interference 215

WILLIAM F. BATTIG

Communalities of Transfer with Other Research Areas 217
Analysis of Component Processes in Intertask Transfer 220
Intratask Facilitation and Interference 226
Implications for Motor-Skills Transfer Research 233
Summary and Conclusions 238
References 240

## Comments on Professor Battig's Paper 245

PAUL W. FOX

References 253

## *Chapter 6.* Information Feedback 255

INA McD. BILODEAU

Labels and Definitions 255
Hypothetical Properties of Feedback 257
An Area Delimited 259
Temporal Locus 260
Frequency of IF 269
Augmented IF 274
IF Transformations 279
References 289

## Comments on Professor I. McD. Bilodeau's Paper 297

L. E. BOURNE, JR.

Some General Issues 298
The Function of IF 299

Temporal Locus .......... 302
Other Variables .......... 308
Concluding Remarks .......... 311
References .......... 312

*Chapter 7.* **Retention** 315

EDWARD A. BILODEAU

Controlled Recall .......... 333
References .......... 347

Comments on Professor E. A. Bilodeau's Paper 351

HARRY P. BAHRICK

References .......... 359

*Chapter 8.* **Tracking Behavior** 361

E. C. POULTON

Introduction .......... 361
The Track .......... 364
The Display .......... 369
The Control .......... 380
The Control System .......... 384
The Man .......... 389
The Training Procedure .......... 398
References .......... 404

Comments on Dr. Poulton's Paper 411

GEORGE E. BRIGGS

Display Variables .......... 411
The Control System .......... 416
The Human Operator .......... 419
Training Procedures .......... 421
Concluding Remarks .......... 423
References .......... 423

*Chapter 9.* **Cybernetic Theory and Analysis of Learning** 425

KARL U. SMITH

Cybernetic Approach to Learning .......... 426
Learning Feedback Control of Basic Movements .......... 431
Delayed Feedback and Learning .......... 436
Cybernetic Analysis of the Spatial Factors in Learning .......... 446
Receptor Cybernetics and Perceptual Learning .......... 459
Cybernetic Design and Instrumental Learning .......... 464
Cybernetic Theory in Learning Science .......... 469
References .......... 478

Comments on Professor Smith's Paper 483

FRED ATTNEAVE

*Chapter 10.* **Motor-Skills Learning and Verbal Learning: Some Observations** 489

BENTON J. UNDERWOOD

Unit of Analysis .......... 492
Constituent Analysis .......... 495
Synthesis .......... 498
Theory .......... 504
Retention .......... 509
Some Final Observations .......... 514
References .......... 515

AUTHOR INDEX .......... 517
SUBJECT INDEX .......... 529

# ACQUISITION OF SKILL

# Chapter 1

# A Brief History of Research on the Acquisition of Skill

ARTHUR L. IRION
*Tulane University*

## Introduction

According to tradition, an historical introduction should begin with a definition of the subject matter area to be covered and the identification of some individual to whom can be assigned the credit or blame for first having worked in that area. Thus, in the field of rote learning, the historian could define the field by pointing to the experimental situations involving serial and paired-associate learning, the characteristic verbal materials usually employed, and the traditional problems that have been investigated. According to his preference, he could identify a founder by referring to Aristotle's views on association, to one or more of the British associationists, or to the experimental crystallization of some of these ideas by Ebbinghaus. However, in the field of motor learning and motor behavior, the definition of the field is by no means as easy to accomplish. Moreover, there does not appear to be any one individual who can legitimately wear the mantle of The Founder. Pioneers there were, and aplenty, but a claim to foundership by any one of them seems somewhat hollow. As often as Bryan and Harter's (1897, 1899) curves for the learning of telegraphic skill have been reproduced in elementary textbooks, it would be difficult to attribute to these studies an importance equal to the findings of Pavlov in the field of conditioning, and neither can the early work of Swift (1904, 1905, 1906, 1910; Swift & Schuyler, 1907) really be compared with the systematic investigations of Ebbinghaus. The work of the would-be founders, in such comparisons, appears to be limited in scope as well as in imaginativeness. The field of motor skills lacks a pioneer who possessed a high degree of creativity, who engaged in a wide variety of researches, and who promulgated a number of strong opinions that could come to be accepted as dogma by subsequent workers in the field.

While it is undeniably true that here, as in other situations, want of identifiable ancestry can be an embarrassment, it is also true that the field of motor skills probably owes some of its vitality to just that deficiency. A founder tends to define the field of discourse, to set its methods and its problems, and, all too often, to bequeath a legacy of sterility as well. Too frequently, the dead hand of the founder continues to rest on the helm of the ship and the weight of the tradition he has established to limit the scope and imaginativeness of the work of those who follow. The field of motor skills does not suffer from a lack of variety of approach. Indeed, the approaches and methods are so extremely various that there is some difficulty in defining, in a sensible way, what the field of motor skills is.

It is undoubtedly true that those who work in the skills area develop an intuitive feeling for what is, and what is not, properly to be considered a part of the skills field. Were 10 such workers to be assigned the task of dichotomizing the literature of experimental psychology into skills research as opposed to non-skills research, there would undoubtedly be a high degree of similarity among the resulting assortments. However, it could also be predicted with almost equal certainty that it would be very difficult to infer a rational basis for their classifications. There are many diverse skeins of interest and activity in the motor-skills area and most of these overlap, to a considerable extent, with the whole remaining field of experimental psychology. So significant is this overlap, that an excellent argument could be made for the proposition that the history of motor skills research *is*, in large measure, a history of experimental psychology. The traditional areas of learning, sensation, perception, cognition, work phenomena, physiological correlates of behavior, motivation, and psychometrics have all been represented in the research on skills and, while not all of the significant work in these areas has had an immediate relevance to the research in skills, it is difficult to think of an important research finding in the field of motor skills that has not made a contribution to our understanding of one or more of the basic problem areas.

While part of the difficulty of defining the area of motor skills may stem from the previously mentioned lack of a founding father, the lack of a serviceable basic concept is probably a more significant deficiency. Other areas of investigation have been more fortunate in this respect. The concept of the *idea*, developed so thoroughly by the associationists, proved to have great value for the workers in the field of rote learning. The concept that was the obvious analog to the idea and that should have been useful to the workers in the field of motor skills, the Cartesian reflex, turned out to be not so very useful, after all, in the skills context. Quite evidently, the concept of the reflex has been much more useful to the physiological psychologists and the Pavlovians than it has been to the workers in the area of skills. The difficulty, of course, is that the idea and the reflex are excellent concepts for describing behaviors that naturally break into discrete units (or that can be broken into such units without the introduction of too great a degree of artificiality). The concepts are not nearly as satisfactory for describing behaviors that are essentially continuous. It is true, of course, that some motor performances are better described than others as a series of discrete events, and it is also true that the exigencies of the laboratory have forced workers to consider motor events in a conceptual framework of discrete events whether or not the resulting approximation to reality is particularly apposite. While it cannot be denied that because of the former circumstance, and in spite of the latter one, considerable advances have been made in the understanding of motor skills in terms of discrete events, it is also possible that the use of a model for discrete events may, in the long run, prove to be more harmful than beneficial.

In their excellent review, E. A. Bilodeau and Bilodeau (1961) offer a "deliberately naïve" definition that differentiates the motor-skills field from other areas in terms of the relative importance of the hand, the tongue, and the eye; and, which, with the further restrictions they place upon it, turns out to be more deliberate than naïve. Perhaps, and for definitional purposes only, two dimensions—one dealing with response availability and the other with response selection—can be offered to supplement their definition.

For those who have been concerned with the *learning* of motor skills, at any rate, it would seem that there has been much more concern with the problem of acquiring responses, as such, than with the problem of selecting from among responses that have been acquired previously. Learning situations in which response availability is low, and in which learning consists of the acquisition of appropriate responses, include such commonplace situations as are involved in learning to swim, to play the piano, to drive a golf ball, and (in an earlier and less automated era) to fly an airplane. In apparatus terms, the dimension is one that extends, for example, from the pursuit rotor (Koerth, 1922), through the Complex Coordinator (Melton, 1947a), to the simple reaction-time situation.

The other dimension is the dimension of response selection. Again, in terms of experimental situations, the dimension stretches from the simple reaction-time situation at the low end of the selection dimension to, let us say, performance on the Selective Mathometer (Noble & Farese, 1955) at the high end.

While it would be of dubious value to attempt to locate tasks along such ill-defined dimensions for purposes of making these positions independent variables for research purposes, the dimensions may be useful, nevertheless, in helping to define what the field of motor skills is. More usually than not, in both practical learning situations and in the laboratory, a particular learning task falls at some intermediate point along both dimensions. Tasks that are high in response availability and low in response selectivity tend to be trivial from the learning point of view and are not much studied (the reaction-time experiments having another purpose than the study of the acquisition of skill). Learning situations that involve a high amount of selection from among very unavailable responses are not much used in the laboratory since they are extremely difficult. As an ultimate example of this combination, one can contemplate the number of hours or "trials" of practice that would be required to enable a completely naïve subject to play one of the more difficult Beethoven piano sonatas precisely and from memory. In the research situations that have been used, either response availability is low and selectivity is also low (as in the case of the rotor) or response availability is high and selectivity is high (as in the case of the Selective Mathometer), or moderate degrees of selectivity and availability are combined in the task (as in the case of the Complex Coordinator). For the purpose of defining the field of motor learning, it can be noted that workers generally have dealt with experimental situations that could be located toward the low end of the response availability dimension and in combination with low to moderate degrees of response selectivity.

The history of research in the motor-skills area is marked by several noticeable quantitative and qualitative trends. In quantitative terms, the sheer total output of research papers, plotted against time, shows that, from relatively modest beginnings at the very end of the 19th century, there was a steady and positively accelerated growth in the number of papers published per year until a peak was reached sometime toward the end of the decade 1950–1959. Since that time, there has been a steady decrease in total research output. In part, the recent decrease can be accounted for by the fact that some of the previously very active contributors to the skills field have now directed their research on other courses while few active, new contributors are appearing on the scene to take their places. Partly, too, the decrease results from the lowered amount of federal support (in the United States) for investigations in the skills area. There is also the possibility that some of the decline is attributable to the fact that one or two of the major research issues of the 1950's have been beaten almost to death in the laboratory. Finally, there is more than a suspicion that the decrease is, in part, illusory—that it results from an error in classifying research. In a rapidly changing field of investigation, active research workers may well be drawn into bypaths of research that, for them, represent the central part of their problem, but which are counted as something else by the historian who is always more inclined to be bound by the traditions and classifications of the past. Even if this last source of error is taken into subjective account, however, it would seem that there has been a decrease in the volume of work that can be classified as skills research, and that this decline began within the past decade.

In general, the history of skills research can be broken into two rather obvious segments, and perhaps into some less obvious ones as well. The major break in skills research appears to have occurred between 1943 and 1949. The lower limiting date is taken as the date of publication of Hull's *Principles of Behavior*. Although the first important fruits of Hull's theory did not begin to appear in the motor-skills literature until the late 1940's (Ammons, 1947a, b; Kimble & Horenstein, 1948; Kimble, 1949), there can be no doubt that the *Principles* exerted a decisive influence on the field. The upper limiting date, which is more local and more arbitrary, is taken as the approximate date of the emergence of the postwar research program of the United States Air Force that supported work in the skills area so heavily and so effectively. If a definite date has to be assigned to the breaking point, 1945 would appear to be an appropriate one since it coincides with the end of World War II.

The causes of the change are multiple. In addition to the emergence of Hull's theory, and other theoretical viewpoints, and the postwar support of research, there existed, at the close of World War II a considerable number of young experimental psychologists who had been exposed to skills problems in the research and aircrew selection programs of the Army Air Corps during the war. (In Great Britian, a similar influence was exerted by the program sponsored by the Royal Air Force and directed by Sir Frederic Bartlett.) At the same time, new and much improved techniques for data analysis came into general use. As might be expected, these influences affected different individuals in different ways, but of the strength of their combined impact upon the research effort of the 1950's, there can be no doubt. Almost without exception, the major contributors to the literature of this decade were either (1) operating under a Hullian or a Hullian-derived theoretical framework, or (2) supported in their research efforts by the USAF (or the RAF), or (3) former military or civilian workers in the aeronautical programs of World War II. Many of the prominent research workers of the 1950's reflected the influence of at least two of these variables and not a few were influenced by all three.

If there can be some dispute concerning the forces that produced the change in motor-skills research that occurred during the period in question, there can be no question as to the profound character of the change itself. In purely quantitative terms, this was the change that was accompanied and followed by the great surge of research that reached its peak some 10 or 15 years later; but the change was much more than a quantitative one. More was involved, also, than the selection of more and better topics for research, although this occurred too. At a more fundamental level, the conception of experiments, the efficiency of their designs, and the deftness of their execution underwent a basic revolution. There were, of course, some excellent pieces of research before 1945, although the average quality of the pre-1945 research seems woefully inadequate as measured against the technical standards of today. A part of the increase in the power and effectiveness of research in skills during the postwar period was a result of the greater sophistication that the later experimenters possessed in terms of variables, theoretical structures, and the practical business of experimentation. However, perhaps the most profound difference was brought about by events that occurred outside of the skills area—the development of new statistical techniques for handling skills data. One who gives more than a cursory glance at the pre-1945 literature in skills is immediately struck by the fact that the investigators seemed to have been handcuffed by their inability to treat their data in meaningful ways. Occasionally quantitative data were presented in the form of means with their probable errors and/or with appended critical ratios for pairs of conditions. More often, the investigator presented his data in the graphic form of learning curves. The great preoccupation that the early experimenters exhibited with respect to learning curves probably reflects, as much as anything else, the want of a technique for treating their data in any other meaningful way. The lack of analytical power also had another influence on prewar research. Since the research worker was more or less limited to the drawing of one or more curves and letting the reader come to his own conclusions, the experimenter was virtually forced to confine his investigations to variables that produced large effects. Lacking the tools for analyzing even large differences, he did not dare to venture into the investigation of variables that produced more subtle effects. In this instance, tradition compounded the difficulty. Early experimentation in all areas of psychology tended to use very small numbers of subjects. The use of two or three hundred subjects in each experimental group would allow an investigator to make some shrewd and valid decisions about subtle effects even if he did not have sophisticated statistical tools at his disposal. It is an ironic fact that the tools for powerful analysis and the tradition of using large numbers of subjects arrived on the scene at the same time. While no one is apt to deplore the use of respectably large groups in combination with refined tools of statistical analysis, it is unfortunate that the large-group alternative to statistical elegance was not more employed in earlier times.

If the major break in the character of motor-skills research is accepted as having occurred, and approximately at the specified time, the further subdivision of the first of the two major historical eras may be suggested. This suggestion is advanced in a much more tentative fashion than the suggestion regarding the major division, both as to its reality and as to the date on which the divisional line should be drawn. However, it does appear that the pre-1945 period exhibits a shift or a change of emphasis that took place between 1925 and 1935. Before 1925, much of the work appears to have been exploratory in character—at least as it is examined in retrospect. If a definite date is desired for the end of the early period, perhaps the appearance of McGeoch's (1927) review of the literature on skills could serve as an arbitrary reference date.

We are left, then, with three periods. The first of these extends from the 1890's to about 1927. The second lies between 1927 and 1945. The final period extends from about 1945 to the present time.[1] The first period was a period of definition and exploration. The second period was marked by a considerable increase in the sophistication of experimental work, by the borrowing of techniques and concepts from other areas of psychology, and particularly, by the emergence of several theoretical formulations that gave structure to the work of the period. It was during the second period that many of the important ideas and concepts that were to produce such a large harvest of research during the 1950's began to emerge. Finally, a few general remarks should be made about the third period since it was during this era that more than half of all of the research that has ever been done in the skills field was published. Indeed, the third period is more characterized by this surge of publication than by anything else. Such an explosion of research could probably only have occurred under just the proper set of circumstances which, in this case, would seem to include the following ones. (1) World War II had produced a large and fairly homogeneous group of young psychologists who had gained some sophistication and interest in the skills area. Many of these psychologists had accumulated a backlog of research ideas during the war years that could not be put to the experimental test until after the war. (2) By the end of World War II, several attractive theoretical models were in existence that could give orientation and direction to the work in skills. (3) Techniques of data analysis that had been developed during the 1930's and 1940's, and were much more sophisticated and powerful than the techniques used by prewar psychologists, were ready for use in psychological research. (4) There was generous financial support for research in the skills area. Under these circumstances, it was almost inevitable that the rate of production of research should have increased, which it did, by a factor of approximately 3.

The contribution of each era to the total literature of the skills field can be estimated only approximately, since the estimates must reflect someone's judgment as to what should be, and should not be, counted as skills research. For what it may be worth, however, such an estimate is offered. It would appear that the first era contributed about 8% of the total published research in the skills field; the second era contributed about 17%; and the third era contributed nearly 75% of the total amount of published research. These percentages, as noted above, are only approximations and, of course, they do not reflect contributions to knowledge—only contributions to the literature as measured by number of papers published.

---

[1] It is possible that it would also be profitable to fractionate the third major period, the one that extends from 1945 to the present time. Some shift of emphasis does appear to have occurred late in the 1950's. However, if such a natural division exists, it is obscured by the fact that we are still in the period of transition.

A few remarks should also be offered concerning the phenomenon of productivity as it applies to research in the skills area. In most fields of research, when number of authors is plotted against number of papers per author, a J-curve is obtained which symbolizes the fact that most authors are extremely unproductive while an exceedingly few authors are very productive. The same situation exists in the field of motor skills. Nearly half of the "contributors" to the skills field have been the co-author of one paper only. About one author in a hundred has contributed as many as 20 papers to the literature, and an author who has written more than three papers is in the top 10% on the scale of productivity. It may be that this distribution of productivity results from the fact that some research workers tend to scatter their efforts over a number of fields while others tend to concentrate in one area. However, since the same general curve seems to be obtained for contributors in other fields and, especially, since contributors to the *Journal of Experimental Psychology* (regardless of field) are distributed in the same way, it would seem to be more likely that the J-curve reflects a basic fact concerning general scientific productivity, at least as far as psychologists are concerned.

A few concluding comments should be made concerning the systematic predilections of those who have worked in the field of motor skills. Through the years, this group has been remarkable for its homogeneity of systematic viewpoint and, during the most recent period, for the similarity of the age and background of its members. Earlier workers in the skills field were mostly quantitatively minded functionalists. During more recent times, they have been behaviorists. This is not to say that a variety of theoretical viewpoints have not been represented. It is to say that there has been very little controversy concerning ultimate aims and purposes. The earlier work is remarkably free from the methods and strictures of structuralism. In later years, the intrusions of gestalt theory have been so infrequent as to make each instance remarkable. There appears to be no example of psychoanalytically oriented research in the skills literature, and it is difficult even to conceive of the role an existentialist might play in the psychomotor laboratory. Homogeneity of viewpoint can, no doubt, be a curse as well as a blessing, and it is not within the province of this chapter to speculate about the kinds of work that might have been accomplished under different systematic leadership, but history must record that those who have worked in the skills field have been behaviorists.

## The Early Period[2]

One of the first areas in the skills field to receive research attention was the problem of distribution of practice. Over the years, more motor-learning studies have been conducted with respect to distribution phenomena (and the related problem of reminiscence) than have been concerned with any other specific area of investigation. In the beginning, interest in the problem of distribution was stimulated by research findings in other learning situations (for the problem can be traced back to Ebbinghaus), and such early workers as Jost (1897) had not only demonstrated the relatively greater efficiency of distributed over massed trials but had also engaged in theoretical speculations that were relevant to the distribution problem. From such beginnings, the research in the area of distribution of practice appears to have gone through three periods which coincide, roughly, with the three periods into which the whole history of skills research has been divided.

Early work, such as that by Starch (1912) and Pyle (1913, 1914), was largely concerned with determining whether or not the introduction of rest periods between practice sessions did, in fact, affect performance. The early work was also marked by a number of experimental preferences, two of the most important of which were the following:

1. There appeared to be a considerable amount of interest in generating

---

[2] Detailed histories of several important areas and topics of research are not covered in this chapter. In general, areas that are important for learning and the conditions of practice have been covered. Not covered is all of the work on the mechanics of movement, the discrimination of stimulus situations, quantitative set, the dynamogenic effects of muscular tension, the psychological refractory period, and application of information theory. Some of this work has been reviewed by Adams (1961a, 1962, 1964), Attneave (1959), Courts (1942), Meyer (1953), Poulton (1957), Saufley and Bilodeau (1963), and Smith and Smith (1962). The reader is also referred to the ensuing chapters of this volume, particularly those by Adams, Poulton, and Smith. It is a matter of particular regret that the excellent and extremely comprehensive work of Fitts and his associates in the broadly defined area of human engineering cannot be covered here. The interested reader should consult the following works (which are barely representative of the enormous volume of material): Fitts (1951, 1958), Melton and Briggs (1960), Bahrick, Fitts, and Briggs (1957).

a universal rule concerning distribution effects so that optimal work-rest schedules could be devised. There was a concomitant tendency to attempt to transpose results obtained under widely different experimental conditions. Since not all situations yield the same distribution benefits, and some tend to show an advantage for massed trials, some confusion was introduced into early attempts to summarize the findings.

2. Early experiments tended to confound length of practice periods with length of the interpolated rest periods. This tendency reflects a preference for which variables should be confounded for it must be remembered that there are three basic variables in distribution-of-practice experiments: amount of practice, length of rest, and total elapsed time. It is not possible to conduct an experiment in which two of the variables are not confounded. Early experimenters evidently felt that it was important to control the total amount of elapsed time while keeping total amount of rest constant. Under these circumstances, as amount of practice per session increases, length of rests between sessions also increases.

In very general terms, the findings of the early research on the problem of distribution of practice may be summarized as follows:

1. In most learning situations, but not in all of them, distributed practice is superior to massed practice when performance is plotted against some measure of amount of practice.

2. In general, long rests are superior to short rests, although there are obvious limitations to the practical implementation of this finding.

3. As a rule, short practice sessions between rests yield better scores than do long practice sessions.

Much of the early work on the distribution of practice is summarized in the review of the literature by Ruch (1928). The problem of fatigue, as it may relate to the phenomena of distribution, was summarized by Robinson (1934).

A second area of early research, and one that, in its original form, was more or less confined to the early period, was the whole-part problem. The whole-part problem seems to have represented many things to many people, but the one thing it does not seem to have been is a problem, in and of itself.

The difficulty is that the whole-part problem overlaps with a considerable number of other problems. For example, it tends to merge with the problem of distribution of practice in that each of the parts may be considered to be a shorter unit of practice than is the practicing of the whole task. If the parts be practiced in alternation or succession, practice on each part is distributed by the length of time required for practice on the other part or parts. However, the fact that the parts must be practiced in some order also throws the whole-part problem into the category of transfer-of-training experiments including designs for the study of proactive and retroactive inhibition. If a task, *X*, is broken down into parts *A* and *B* (so that $A + B = X$), if *A* is practiced first, *B* is practiced second, and finally a test is made of $A + B$, negative transfer from *A* to *B* can be considered as proactive inhibition, and negative transfer from *B* to *A* (in $A + B$) can be considered as retroactive inhibition. Moreover, it should be noted that, although the early experiments on the whole-part problem were not designed to test for combination effects, these investigations can be considered in this way. If *A* and *B* are separate components of *X* and if there is a combining rule for predicting the value of *X* if values are known for *A* and *B* separately, the whole-part problem becomes a special case of the combination problem. Finally, it should be noted that the whole-part problem is a part of the problem that deals with length-difficulty relationships, and it is out of this context that the whole-part problem seems to have arisen. The length-difficulty relationship, itself, has been studied in conjunction with a number of other variables. The classical paper on the length-difficulty problem by Lyon (1914) also dealt with this relationship under conditions of massing and distribution and reported a strong interaction effect, the longer materials being disproportionately easier to learn under distributed than under massed practice conditions. In the case of the whole-part problem and its relation to research on the distribution of practice, there have been experiments in which both whole and part methods of practice were used with and without the introduction of additional rests. For example, Crafts (1930) used a letter-number substitution task and found that the whole method was about equal (or slightly inferior) to the part methods he employed when practice was massed, but that the whole method was distinctly and consistently superior under conditions of distributed practice.[3]

As far as other areas of investigation were concerned, the first historical period was one in which fundamental concepts were being formed and explorations were taking place, but frequently these did not have any particular relevance to the skills field at the time they occurred. For example, the classical papers on transfer of training by Thorndike and Woodworth (1901) were later to exert a considerable influence on the work in skills, but they were not oriented in this direction at the time they were written. In a somewhat similar way, although in these cases, skill-learning was involved, the exploratory studies of Bryan and Harter (1897, 1899) and Starch (1910) on learning curves, and the studies of Swift (1904, 1905, 1906, 1910; Swift & Schuyler, 1907) on the learning and retention of skills, served mainly to define problem areas for later investigators.

---

[3] An interesting "solution" to the whole-part problem has recently been offered by Annett and Kay (1956). They point out that a series of recurring events is probably best learned by the whole method. If, however, the responses of the subject serve to alter the series of events, part practice will tend to reduce the variability or randomness of the series. They pose the paradox that "if the elements of a task are highly independent the task is best learned as a whole, but where the elements are highly interdependent they should be split up and the task learned in parts."

## The Second Period

If the first period was largely devoted to exploratory work, the second period was marked by the emergence of a number of theoretical points of view and by the sharpening of concepts. Also toward the end of the second era, many of the problems of methodology and data analysis began to come into focus. To a considerable extent, the second era is distinguished by the borrowing of concepts and methods from other areas of psychology. In addition, several new areas of skills research became established during this period. If, in retrospect, some of the theorizing and experimental work appears to be somewhat naïve and, perhaps, crude, it must be remembered that, without the preparatory work of this era, the great explosion of fruitful research that followed World War II would not have been possible.

During the second era, work on problems of distribution of practice continued to be published, and at an increasing rate. To a considerable extent, by the advent of the 1930's, the use of the concept of fatigue as an explanation of distribution effects had fallen from favor and a considerable number of other special theories and hypotheses had been advanced. Only two of these seem to have had much impact on the skills field. These were (1) the two hypotheses concerning age maturation and stimulation maturation advanced by Wheeler and Perkins (1932), and (2) the Snoddy hypothesis (1935).

Both of the Wheeler and Perkins hypotheses tend to identify the process of learning with the process of maturation and, in effect, to consider learning as a special form of stimulus-induced growth. The age-maturation hypothesis merely holds that the ability to benefit from practice depends upon the age-maturational level of the subject. Doré and Hilgard (1937) describe the stimulation-maturation hypothesis, as follows: "At any given age-maturational level, the optimum rate of stimulation to secure improvement in a given function depends upon the normal rate of stimulation-induced maturation in that function" (p. 246). An implication of this formulation is that, given a particular age-maturational level, stimulation (practice) will give rise to a rate of growth (improvement) that is related more closely to total elapsed time since the beginning of practice than to the total amount of practice given in that time. This last implication of the Wheeler and Perkins theory was specifically tested by Doré and Hilgard (1937) in a skills situation. Using the pursuit rotor, they distributed different amounts of practice through the same total amount of elapsed time. Their acquisition curves, when performance was plotted against trials in the conventional manner, showed the usual distribution benefit. When performance was plotted against elapsed time, however, the improvement curves tended to coincide, thereby lending support to the theory. The results obtained by Hilgard and Smith (1942) also seem to lend support to the hypothesis, although later findings by Bell (1942) do not. McGeoch and Irion (1952) have attempted to reconcile these contradictory results by pointing out that, in view of what is known about the relative effects of practice and rest, it is possible to predict, on purely empirical grounds, that a range of practice-rest conditions exists for which the predictions of Wheeler and Perkins should be fulfilled.

The Snoddy hypothesis posits two opposed processes of mental growth called primary growth and secondary growth, respectively. Primary growth is supposed to occur early in practice and to be a joint function of amount of practice and amount of interpolated time. The cumulative mean of the scores during practice is held to be a measure of the amount of primary growth, although this measure has been attacked on logical grounds by Doré and Hilgard (1938) and on empirical grounds by Payne (1942). Secondary growth is supposed to be a function of the base established by primary growth. It tends, therefore, to occur later in practice. Secondary growth is held to be unstable in the sense that it dissipates spontaneously in time and, hence, benefits from massed trials.[4]

Snoddy's results (1935), and those of Humphreys (1937), appear to agree with the predictions that can be made from the theory, as do the findings of Renshaw and Schwarzbek (1938). However, Bell's (1942) findings offered small support to the theory, and the results obtained by Doré and Hilgard (1938) were nearly devastating to it. Not only were the findings of Doré and Hilgard contrary to a clear prediction from Snoddy's theory, but their replotting of earlier data seemed to change the interpretation of them in a radical manner. More significantly, however, Doré and Hilgard offered a two-factor theory of their own based on (1) learning and (2) work decrement. Their view, in its more sophisticated contemporary form, is reasonably acceptable doctrine today.

During the second period, investigations of the reminiscence phenomenon in motor-learning situations had not really begun to occur. The interested reader should consult Buxton's (1943) review of the literature and Ammon's (1947a) theoretical paper.

---

[4] It was also held that individual variations tended to decrease in primary growth and to be enhanced in secondary growth. Although it is probably not possible to design an unequivocal test of this implication of Snoddy's theory, it would seem that, if a single long rest were to be introduced late in a series of massed trials, variance should be greater on the immediate prerest than on the immediate postrest trial.

A second area of research that emerged during the second period, and one that was destined to grow in importance during the subsequent era, dealt with the problem of transfer of training. The early findings of Thorndike and Woodworth (1901) served as a basis for several early studies of transfer such as those by Pyle (1919), McGinnis (1929), and Crafts (1935). These studies, which were exploratory in character, yielded a finding that was to become discouragingly familiar to later investigations: that it is much easier to produce positive than negative transfer effects in motor-learning situations. During the second period of our history, much of the work on transfer appears to have been influenced by the formulation of the Skaggs–Robinson hypothesis (Robinson, 1927), by Bruce's (1933) formulation of the conditions of transfer, and by the growing literature on proactive and retroactive inhibition in the field of verbal learning (reviewed by Britt in 1935). These influences are revealed in the studies that attempted to produce interference effects (Buxton & Henry, 1939; Buxton & Grant, 1939; Buxton, 1940). Attempts to produce retroactive inhibition failed to yield absolute losses under negative transfer conditions although "relative" retroactive inhibition was obtained. In retrospect, a study by Siipola and Israel (1933) appears to have particular value in that it showed that amount of interference between two tasks depends, in large measure, upon the relative amounts of training on the two tasks. The apparatus used by Siipola (1940) also appears to have been the prototypical inspiration for a number of laboratory devices used by later experimenters.

A somewhat different theoretical background is reflected in the two well-known studies by Hovland and Sears (1938; Sears and Hovland, 1941). These studies were based on a model of conflict behavior, but they can be considered in the context of classical transfer theory. Their studies are of particular value in that, in addition to measuring "amount of interference," their technique permitted an analysis of the kinds of interference or conflict resolutions that were taking place. Their finding that "the probability of blockage as a reaction to conflict increases with the approach of the strengths of the conflicting responses to equality" is in essential agreement with the earlier, and less analytic, finding of Siipola and Israel (1933) that amount of interference is maximal when the two interfering habits are of approximately equal strength.

Finally, it should be noted in passing that two important studies of bilateral transfer of training, one by Bray (1928) and the other by Munn (1932), were conducted during this second historical era.

The thread of psychometrics runs through the entire history of motor-skills research. From the beginning, no doubt, skills situations have often been perceived as likely situations for the measurement of specific aptitudes or for predicting particular kinds of performance. The heaviest use of skills situations for such purposes occurred during World War II in connection with various aircrew selection programs, but the general notion of psychomotor testing is much more ancient than this. In fact, Whipple's *Manual* (1910) contains descriptions of a number of simple psychomotor tests, several of which were still being used for aircrew selection purposes during World War II.

In the context of this chapter, the ability to predict success in specific occupations, military assignments, or training situations by means of psychomotor devices and apparatus tests of the kind described by Melton (1947a, b) is trivial. This does not mean, of course, that such prediction is a matter of small consequence in other contexts. It should also be noted that programs aimed at personnel selection at the most practical levels of application have frequently yielded ancillary benefits to the more fundamental areas of skills research. Thus, a considerable part of the postwar research on skills was conducted with equipment, and in situations, that had originally been designed primarily for personnel selection purposes. Many of the qualities of a satisfactory selection test such as objectivity of scoring, ease of administration, reliability of scores, and the existence of a wide range of individual variation in scores are also qualities that may be desirable in research instruments.

However, there is another aspect of the psychometric trend in relation to skills research that is not at all trivial in the context of the present review. This aspect concerns the study of individual variation in performance, the use of test scores to determine the factors entering into skilled performances, and, especially, the study of the course of individual variations and factor structures under the conditions of practice.

During the earlier portion of the second period, research in the psychometric area tended to be concerned with the measurement of the dispersions of scores about mean performances and with the contributions to total variability of such factors as age and sex. There were also a number of studies that concerned themselves with the development of motor skills and abilities in childhood and adolescence, of which the studies by Hicks (1931), Yarmolenko (1933), McCloy (1935), and Espenschade (1940) may be taken as representative. Some attention was also directed toward the intercorrelations of scores among different skill-testing situations (the general finding being that the intercorrelations were positive and low), and the number of such studies increased as the technique of factor analysis came into use (Goodenough & Smart, 1935; Campbell, 1934; Buxton & Humphreys, 1935; Buxton, 1938; Seashore, Buxton, & McCollom, 1940). The second period ended with the publication of a stunning series of papers by Woodrow (1938a, b, c, 1939a, b, 1940, 1946). Strangely enough, these papers, which should have marked a major turning point in the research on skills (and, indeed, on learning research in general) appear to have had little impact in their own time and, in fact, their implications have not yet been fully explored. The immediate impact of Woodrow's work on the field of motor skills was exerted through the Ph.D dissertation (and a subsequent publication) of one of his students, Mary J. Kientzle, whose two papers (1946, 1949) were, as will be seen, quite influential in the area of research on distribution of practice during the third epoch of our history.

It was Woodrow who first exploited the interesting pattern of correlations that has frequently been observed in learning situations. The phenomenon is a regularly appearing one. When successive trial scores are correlated, the correlations, between the initial score and the successively following scores that are generated as practice progresses show a marked tendency to decrease; that is, score on trial 1 will usually correlate more highly with score on trial 2 than with the score on any subsequent trial. Similarly, the correlation between scores on trial 1 and the score on the *n*th trial will typically be the lowest of the set. At the same time, there is a less pronounced, but very consistent, tendency for the correlations between successive trials to increase in magnitude from the beginning of practice to the end of practice. The first explanation of these findings was offered by Woodrow (1938b) who noted:

> "Now suppose, for example, that final performance depended mainly on ability Y and initial ability mainly on ability X, and that individual differences were greater in ability Y than in ability X—then individual differences would tend to increase with practice; and vice versa.
>
> "The change with practice in the pattern of abilities brought into play in the execution of the task also could account for the falling-off with practice in the correlation coefficient, since the abilities here supposed are either uncorrelated or correlated only to a small degree . . . Now when the scores are accounted for in terms of uncorrelated abilities, it is obvious that if initial scores were altogether dependent on ability X and final scores altogether dependent on ability Y, the correlation between initial and final score would be zero; and further, that any considerable change in the degree to which the scores depended upon the two uncorrelated abilities would result in a correlation between initial and final score markedly below a perfect correlation. The greater the change in the ability pattern, when the abilities are uncorrelated, the lower should be the correlation between initial and final score" (p. 277).

Several of Woodrow's conclusions and observations have a remarkably contemporary aura about them. Writing in 1938, Woodrow noted that marked changes in factor loadings may occur as a result of practice (1938a). He concluded:

> "If the goodness of the scores be regarded as determined by a *set* of co-operating but independently variable abilities, then practice may be regarded as a change in the conditions under which the various constituents of this set of abilities operate" (p. 227).

Again Woodrow noted:

> "If, then, initial and final measures are equally reliable, as is here the case, it is not true that there is any *general* tendency for test-performances to become with practice more dependent upon or better measures of some supposedly common factor, such as "g," intelligence, or speed. One reason why test scores have sometimes been supposed to do so is probably the fact that as a rule previous investigators have not used initial scores of as high reliability as the final scores" (p. 227).

He also noted that gain scores usually correlate highly with final (post-practice) scores, but poorly with initial scores. In 1938(a), and again in 1946, he failed to uncover a general improvement factor. Finally, Woodrow's conclusion (1938b) deserves to be quoted.

> "In conclusion, it is desired to emphasize that the phenomena here discussed form an important aspect of the results of practice and learning. In recent years, experimental psychologists, working largely with the white rat, have made notable progress in the explanation of learning. Applied to practice, however, the theories advanced are helpful mainly in understanding simply the fact that performance improves with practice, i.e., that the practice curves rise. Before any account of the phenomena of practice and learning can be complete, however, it must take account of two other important phenomena; namely, the change in individual differences and the falling-off with practice in the correlation between the first obtained and the last obtained scores" (pp. 277–278).

Several other areas of research also began to receive significant amounts of attention during the second period. The most important of these was the area of retention and forgetting. It had long been assumed that motor habits were better retained than verbal habits, but at the beginning of the second era, this assumption was called into question by McGeoch and Melton (1929), who reviewed the literature on the subject and then compared the retention values of mazes and nonsense syllables, each task being presented at three levels of difficulty. Although their conclusion that retention differences were small was called into question by Freeman and Abernethy (1930, 1932), subsequent findings of McGeoch (1932) and by Waters and Poole (1933) tended to agree with the earlier findings of McGeoch and Melton. Waters and Poole went on to conclude that "the relative retention value of what have been conventionally called motor and verbal habits is a function of the relative degree of learning of the two habits compared." Leavitt and Schlosberg (1944) attempted to equate degree of learning in the pursuit rotor situation and the rote learning situation, and then to compare retention of the two habits. In their experiment, the pursuit rotor habit was retained better than the verbal habit, although the difference was attributable, in part, to the fact that the motor habit exhibited a considerable amount of reminiscence, especially after the shorter retention intervals.

The problem of the relative retention value of verbal and motor habits was not to be raised again in precisely this form and, in retrospect, the issue may be argued from a number of viewpoints. The first of these would have it that such comparisons are meaningless because it is not possible to work with truly comparable situations and that the equation of such factors as degree of learning cannot be accomplished in a meaningful way. A second kind of argument would state that there is little difference between motor forgetting and verbal forgetting because verbal retention is really much greater than it is traditionally supposed to be (cf. Underwood, 1957). A final point of view would concede that motor habits are remarkably resistant to forgetting, and would then marshal arguments to explain this resistance. The most familiar arguments are (1) that motor habits are generally better learned than verbal habits and, hence, are less susceptible to forgetting; (2) that motor habits are less susceptible to interference effects than are verbal habits because the extraexperimental world offers fewer opportunities for the learning of competing habits in the case of motor habits than in the case of verbal habits; and (3) that the situations for measuring motor retention are biased to produce high memory scores while the situations for measuring verbal memory are biased to produce high amounts of interference and forgetting. This third point deserves some amplification.

Forgetting (and interference) effects in verbal learning are usually measured by recall performance. It is probably true that the recall situation in verbal learning experiments is biased in favor of obtaining interference and forgetting. As far as a particular nonsense syllable is concerned (and, of course, a recall score is a composite of performances with respect to particular syllables), an inadequately warmed-up subject is given a 2-second interval in which to make the correct response and is then scored in an all-or-none fashion. Furthermore, in the case of the serial anticipation situation, failure to recall a single syllable early in the series may throw the subject off for the remainder of the recall trial (that is, the recalled events are not independent). By contrast, in motor retention situations, the subject is frequently given a "recall trial" that lasts for 30 seconds or more. During this "trial" he responds to a relatively homogeneous task, may throw off interference effects quickly, regain an appropriate set to respond, and make a quite respectable score. The question as to what the subject is doing during the first 2 seconds of the first relearning period is rarely asked. It is quite possible that, if motor memory were to be scored in terms of a very short recall trial, or if motor memory were to be measured in a situation in which early failure in recall would act to degrade later performance, levels of interference and forgetting in motor-learning and verbal-learning situations might be more alike. It is, perhaps, unnecessary to point out that comparisons of verbal and motor retention can be legitimately made only if the original practice has been highly distributed since, otherwise, forgetting in motor-learning situations would be diminished or obliterated by the occurrence of reminiscence.

## The Third Period

As we have already noted, the third period in our history is marked by an extremely high rate of productivity in all fields of research. Not only were most of the traditional problems subjected to further investigation, but new fields of research activity came into being.

The issue of massed and distributed practice continued to be a dominant theme during the early part of the third period. Much more than half of the extensive literature on this topic may be traced to this era. However, it must also be noted that research in this area came to a rather abrupt halt in the late 1950's, and only a handful of recent studies have been conducted along these lines.

The third period in the history of research on the problem of distributed practice was ushered in by Hull's (1943) introduction ot the concepts of reactive inhibition ($I_R$) and conditioned inhibition (${}_SI_R$) and by Ammons' (1947a) theoretical analysis that was specifically designed to be applicable to the skill-learning situation. The significance of the change consists of more than the introduction of two (or more) new theoretical formulations. It marks the change from thinking about distribution of practice as being a variable that affects learning to a position that views distribution of practice as a variable affecting performance only. As is noted above, this view had already been advanced by Doré and Hilgard (1938).[5]

---

[5] Ammons (1947a) advanced a theoretical analysis that was very similar to Hull's and that was specifically designed to apply to the motor-learning situation in human subjects. Perhaps unfortunately, the general enthusiasm generated by Hull's theory among that segment of psychologists who could become enthusiastic over theories of the general Hullian type tended to carry over into the field of motor-skills research so that Hull's theory was destined to be much more influential in determining future research than was the more specific, and in some ways more satisfactory, theory of Ammons.

At the transition between the middle period and the third period stand two significant studies by Kientzle (1946, 1949), which represent an extension of the Woodrow type of analysis to the phenomena of distributed practice. In addition to the usual finding that distribution of practice results in a higher performance level than does massed practice, Kientzle also found (1) that the experimental groups differing in degree of distribution approach different performance asymptotes, these limits being an increasing, negatively accelerated function of the duration of the rest intervals[6]; (2) that variability increases as a function both of practice and of interpolated rest, so that variability increases more rapidly in the distributed group than in the massed group, and that these increases of variability are independent of the magnitudes of the means; and (3) that correlations between scores on particular trials are dependent on the ordinal numbers of the trials, but not on length of rest. Her conclusion that "the relative contributions of various abilities to variance at any particular trial would be independent of rest," may be regarded as one of the more important statements that have been made about distribution phenomena. The implication of her analysis is quite clear; the effects of rest and the effects of practice represent separate entities.

The design of reminiscence experiments is continuous with the design of distribution-of-practice experiments except that a single rest period is used. The phenomena of distribution experiments and reminiscence experiments also appear to be continuous according to an analysis performed by Buxton (1942), who found reminiscence gains following the first rest periods in a number of previously reported distribution-of-practice experiments. It is probably the proper time to introduce the topic of reminiscence since, with the advent of the third period, three studies of reminiscence in pursuit rotor learning that varied length of the single rest period appeared in quick succession: Ammons (1947b), Kimble and Horenstein (1948), and Irion (1949). Two of these studies, those of Ammons and Irion, also varied amount of prerest practice.

From the standpoint of Hullian theory, the most significant of these experiments was the one by Kimble and Horenstein. They plotted their data for amount of reminiscence as a function of length of rest against Hull's function for the decay of $I_R$ in time, and obtained a tolerably good fit.

In a second experiment, Kimble (1949) used the inverted alphabet task

---

[6] The finding regarding different asymptotes for different degrees of distribution has been corroborated by Adams (1954) but was by no means in agreement with views held in an earlier period as stated, for example, by Lorge (1930). It is probably unnecessary to note that this finding does not imply that distribution enhances learning or that massing degrades learning. The asymptotes apply only to performance under continued practice with the same degree of distribution.

S.C.E.R.T., West Bengal
Date 16. 2. 67
Acc. No. 1972

previously employed by Kientzle. In this experiment, Kimble varied amount of prerest practice and employed, in addition to the standard, massed-practice, control group, a second control group that received highly distributed practice. By this means, Kimble was able to estimate the amount of $I_R$ accumulated after various amounts of prerest practice by measuring the amount of reminiscence against the conventional, massed-practice control group and to estimate the amount of $_sI_R$ accumulated for the same amounts of prerest practice by determining how far short of the distributed practice control group curve the reminiscence points fell. His conclusion was that $I_R$ remained fairly constant during early practice (presumably after rising rapidly to that value in the period before the first rest) and then tended to decrease. The value of accumulated $_sI_R$ rose in a negatively accelerated manner as a function of amount of prerest practice. It is noteworthy, however, that in Kimble's experiment, postrest practice was always given by massed trials, and it is also worth recording that the postrest performance curves of Kimble's subjects failed to yield the warming-up effect that is usually found in situations of this kind. It is possible that the lack of a warming-up effect, in this case, was due to the fact that the 30-second measurement periods were employed and that a warming-up effect might have been observed had shorter periods been used. In each of Kimble's postrest curves, the initial point represents a higher level of performance than the second point; thus it might have happened that the subjects reached a peak of performance somewhere between the beginning of postrest practice and the end of the second postrest trial, by which time they were already well into what Ammons (1947a) has called "the relatively decremental segment of the post-rest performance curve." At any rate, the results obtained by Adams and Reynolds (1954) do not offer much support for the notion that $_sI_R$ operates in postrest performance to depress scores below the level of the distributed practice control group. Using the pursuit rotor, they followed a design very similar to the one used by Kimble, omitting the massed practice control group, but also shifting all groups to distributed practice following the rest. Thus, the experimental groups received various amounts of prerest practice by massed trials, then rested, and finally were given their relearning trials under distributed conditions. All of the experimental groups in relearning seemed to achieve the level of the distributed practice control group, although those experimental groups with the greatest amount of prerest practice approached the level of the distributed practice control group more slowly in relearning than did the experimental groups that had been given the least amount of prerest practice. The conclusion that $_sI_R$ plays a negligible role in motor performance has been arrived at independently by Ammons and Willig (1956) and by Digman (1956, 1959).

Analogous, and corroborative, studies using a manual cranking task have

S.C.E.R.T., West Bengal
Library
1972 67
151
ELL

been performed by E. A. Bilodeau (1952a, b, 1954a). The cranking situation has the virtue of keeping learning factors to a minimum, thereby allowing inhibitory factors to be revealed in relatively "pure" form. Again, there is little evidence that $_{S}I_{R}$ plays a significant role.

The weight of the evidence, then, would seem to suggest (1) that distribution of practice affects performance, but not learning, and (2) that almost all of the distribution or reminiscence benefit can be accounted for in terms of the dissipation of a single inhibitory factor that builds up during practice and dissipates during rest.

To reach these conclusions, of course, does not say very much about the nature of the inhibitory process. A small amount of research, not all of it illuminating, has been performed in respect to this problem. One approach has been to attempt to analyze the nature of the motor performance as it changes during practice and over periods of rest. These studies employ direct observations of behavior rather than inferring performance changes from an analysis of intertrial correlations. A second line of inquiry has had to do with the fact that reminiscence (and, presumably, distribution) gains are found in bilateral transfer situations. A few words should be said about each of these approaches.

The first type of experimentation is exemplified by the studies of Ammons (1951a), Bourne and Archer (1956), and Archer (1958). Ammons employed a kymographic recording technique in the pursuit rotor situation that allowed him to break down time-on-target scores into "number of hits" and "duration of hits." He found that the mean number of "hits" was greater under distributed practice and that, under these conditions, the hits tended to have a longer mean duration.[7] Using a different method of recording, Bourne and Archer came to essentially the same conclusions. They hypothesize that subjects first learn how to get on the target (hits) and later learn how to stay on the target (duration of hits).

The different shape of the acquisition curves for hits and for mean time on target obtained by Ammons (1951a) under massed and distributed practice might suggest that different learning processes were involved under these two conditions of practice, and Bourne and Archer (1956) specifically suggest that one of the results of massing may be the accumulation of inhibition to a degree that interferes with the *learning* of the smooth, circular, following motions essential to efficient performance on the rotor. Archer's (1958) finding led him to hypothesize that inhibitory factors, accumulated during massed practice, produce a competing response (tremors) that interferes with performance and that prevents, to a degree, the learning of the correct responses.

---

[7] Number of hits and duration of hits increased under both massed and distributed practice, but both measures increased more rapidly under distributed trials. It is to be expected, of course, that the finding regarding number of hits could hold true only during the relatively early stages of learning. Perfect performance on the rotor would involve the making of but a single hit per trial, but that hit would last for the total duration of the trial.

However, the conclusion that massing and distribution variables affect learning rather than performance seems to run counter to the findings of a considerable number of investigations and should not be accepted without considerably more supporting evidence than now exists. In particular, it should be remembered that improved performance is bound to be reflected in either a greater number of hits or a greater duration of hits, or both, and that, if total performance is higher under distributed than under massed conditions, this higher performance is bound to be reflected in measures of hits and durations.

The phenomenon of reminiscence in bilateral transfer situations also has a short history. In 1952, Irion and Gustafson (1952) and Kimble (1952) demonstrated that such effects could be obtained. Irion and Gustafson worked from initial performance on the right (preferred) hand to final performance on the left hand while Kimble worked from practice on the left hand to performance on the right. The studies were in substantial agreement in showing that a group receiving a rest at the time of the switchover of hands exhibited a higher subsequent performance than a control group that did not receive such a rest, the groups being compared for performance after the switch of hands.

Grice and Reynolds (1952) followed these studies with one that demonstrated that amount of bilateral reminiscence was a function of the length of the interpolated rest period, a variable known to be important in determining amount of conventional reminiscence. This experiment demonstrated that, while bilateral reminiscence is less in amount than is unilateral reminiscence, it is related to length of rest in approximately the same way. The findings of Grice and Reynolds were corroborated and extended by Rockway (1953), who varied length of rest and amount of first-hand (prerest) practice. Rockway found that performance with the second hand was an increasing function of rest and of amount of first-hand practice.

The findings on bilateral reminiscence are somewhat difficult to interpret. The cogent suggestion has been made by Grice and Reynolds (1952) that "the gains in bilateral transfer indicate the presence of temporary inhibitory factors either associated with generalized postural adjustment or of a perceptual or central character." Early attempts to isolate postural and/or perceptual factors by Goldman (1952) and Duncan (1957) yielded results that were essentially negative as far as the problem of bilateral reminiscence is concerned [but which substantiated earlier findings by Ammons (1951b) and Melton (1947a)]. These studies dealt with the effects of such variables as standing, standing and watching the rotors revolve, standing and watching other subjects practicing on the rotors, etc. However, Adams' (1955) study suggested a line of explanation. He used a rest-filling activity that might be called active or attentive observation. His subjects worked in a conventional (unilateral) reminiscence situation on the pursuit rotor. During the rest period, one group of subjects watched other individuals performing on the rotor. Adams instructed his subjects not only to watch the rotoring performance but also to press a button whenever the individual being observed was "on target." Under these circumstances, active observation (as opposed to ordinary resting) had the effect of depressing the amount of reminiscence obtained. Adams' findings have been extended to the bilateral reminiscence situation in a recent experiment by Rosenquist (1964) that demonstrates that considerable proportions of the bilateral reminiscence effect may be accounted for by active observation during the rest periods.

In the light of the present evidence, it would seem to be a reasonable conclusion that the bilateral reminiscence effect is attributable to the dissipation of inhibition accumulated in connection with response systems that are common to right-handed and left-handed performance. This conclusion, of course, fails to throw additional light upon what the nature of the inhibitory process might be.[8]

A second area of very heavy concentration of work during the third period was transfer of training. Several different lines of investigation were followed. The period began with the appearance of two much-needed pieces of work; the paper by Gagné, Foster, and Crowley (1948) on the measurement of transfer effects, and the theoretical analysis of the conditions of transfer by Osgood (1949).

The investigations of possible interference effects (including proactive and retroactive inhibition) were numerous, dealt with a variety of experimental situations, and tended, in their summative effect, more or less to exhaust the research possibilities in this area. In particular, the work of Lewis and his associates deserves to be mentioned. A very few, representative, studies from this series are cited. Lewis, McAllister, and Adams (1951) used a conventional retroactive inhibition design to study the effects of varying amounts of original training on proactive and retroactive inhibition in the Mashburn Apparatus situation. They found that original learning exerts both facilitating and interfering effects upon the learning of the reversed (interpolated) learning task, although the amounts of these effects did not appear to be closely related to the amount of original learning. Conspicuous amounts of retroactive inhibition were obtained during the early phases of relearning. In a second study, McAllister and Lewis (1951) varied the amount of interpolated learning. Again, retroactive inhibition effects were obtained, amount of inhibition being related to the number of trials of interpolated learning. Somewhat similar findings were reported by Lewis, Smith, and McAllister (1952) for the Two-Hand Coordinator situation.

---

[8] One not very hopeful suggestion concerning this problem which has been offered by Köhler and Fishback (1950) and defended by Duncan (1956) would identify reactive inhibition with such perceptual "satiation" phenomena as figural after-effects. According to this view, the important variable involved in the effects of work is not the work itself, but the kinesthetic stimulation resulting from the work.

Despite these findings, and other like them, it would appear that the research on interference effects has not been productive of very much interference. Except under very carefully arranged conditions, positive transfer appears to be much more likely to occur than negative transfer and, even when optimum conditions for the production of interference are arranged, the negative transfer effects seem to be somewhat weak, quite transitory, and quite apt to shift to facilitation in relearning. Sometimes, as in Duncan's (1953) experiment, conditions that might have been expected to yield negative transfer (on the basis of the analogy with verbal-learning situations) have yielded positive transfer, instead. Massive and persistent negative transfer effects do not seem to be obtainable, probably for the reason that the subject is learning the same thing in the tasks that are supposed to interfere with each other. (Also, as has already been noted, the methodology of motor-learning experiments may be biased against the obtaining of interference effects.) Those who favor an interference theory of forgetting can take some rather meager comfort in the fact that motor skills seem to be impervious to both interference *and* forgetting.

A pair of correlational studies by Lewis, McAllister, and Bechtoldt (1953a, 1953b) supports the notion that one of the reasons for the findings on interference and transfer effects in motor-learning situations may be that the subjects are learning the same structure of skills in both original and interpolated practice sessions. Their findings would seem to be relevant to an observation made by Woodrow (1938a) to the effect that, if two practice situations have a common factor in their gain scores, practice in one should result in positive transfer to the other.

A second approach to the problem of transfer represents an extension of work on the old whole-part problem. At the beginning of the period, Gagné and his associates performed a series of studies concerned with transfer from training on components to performance on the whole task. They also studied the effects of verbal and other kinds of pretraining upon the whole task performance (Gagné & Foster, 1949a,b,; Gagné, Baker, & Foster, 1950; Baker & Wylie, 1950). These investigations generally yielded positive transfer effects. Moreover, the Gagné, Baker, and Foster experiment suggests that, at least as far as pretraining in stimulus discrimination is concerned, pretraining on the more difficult of two useful discriminations produces the greatest amount of positive transfer. In a somewhat more advanced approach to the whole-part problem, rules concerning the combination of component activities in predicting performance on the whole activity have been sought. For example, following Ellson's (1947) demonstration of the practical independence of component performances on the Pedestal Sight Manipulation Test, E. A. Bilodeau and Bilodeau (1954) were able to predict total task means and variances for this test from the use of two components and, using the Stevens Pursuitmeter device, they were able to predict total task means and variances using three scores of component performance. In this study, and in the studies of E. A. Bilodeau (1955a, 1957b), the rule of multiplicative proportions was found to hold approximately true, with the implication that the assumption of independence of component performances is supportable. The first of these studies also indicated that component performances can be independently manipulated [as did the study by E. A. Bilodeau (1956a)], and that total performance, following the multiplicative rule, can be modified accordingly. There results the general hypothesis that, in any complex motor task in which two (or more) component responses must be made simultaneously, total performance may be drastically improved by an increase in the proficiency of the subject on the poorer component, even though that increase may be at the expense of the better component.[9]

A third set of transfer problems revolved around the problem of transfer between tasks differing in difficulty. This is a vexing problem area, partly because the dimension of difficulty is ill defined, partly because many of the studies have combined investigation of transfer effects along the difficulty dimension with study of the effects of other variables such as interference (Barch, 1953; Barch & Lewis, 1954), and partly because these difficulties have yielded contradictory results. The definition of a dimension of difficulty is

---

[9] A somewhat different kind of combination rule has also been proposed by E. A. Bilodeau (1957a) in connection with the prediction of performance in a two-handed tracking task from the performances on similar one-handed tasks. Here, prediction was made in terms of the one-handed scores weighted in terms of the contributions made by each hand to the total performance:

$$t_{R,L} = at_R + (1-a)t_L,$$

where $t_R$ and $t_L$ are the times for traversal using the right and left hands (measured separately) and $a$ and $(1-a)$ represent the input proportions of the right and left hands on the total task. Bilodeau suggests that, with highly practiced subjects, "the time for two-hand performance is equal to the larger term of the equation since prediction can be made on the data from a single hand. Such subjects are said to be coordinated because the time of the hand taking the shorter time is completely distributed over the time of the hand taking the longer time." This work has been extended by E. A. Bilodeau and Ryan (1961).

particularly trying since, in the last analysis, difficulty is a response-inferred concept thereby causing a certain amount of circularity of reasoning to be introduced into studies of response as a function of difficulty. In dealing with rates, for example, Baker, Wylie, and Gagné (1950) obtained more transfer from practice on a fast rate to performance on a slow one than under the opposite arrangement. On the other hand, Ammons, Ammons, and Morgan (1956) obtained greater transfer in working from slow to fast rates than they got in transferring from fast rates to slow ones. Finally, Lincoln and Smith (1951) found mixed results. In other dimensions of difficulty, the general finding seems to be that there is more transfer from a difficult task to an easy one than vice versa (Gagné *et al.*, 1950; Gibbs, 1951; E. I. Jones & Bilodeau, 1952; Andreas, Green, & Spragg, 1954; E. A. Bilodeau & Bilodeau, 1958), but many other studies have reported null results. The problem has been reviewed by Day (1956), who summarizes the findings and discusses some of the methodological difficulties likely to be encountered by those who work in this area.

During the third period, the problem of retention received a considerable amount of attention. Toward the end of the period, that is to say, in very recent times, studies of retention have tended to turn away from tracking-type, time-on-target situations and the concepts these situations use, and have shifted more in the directions that have been taken by studies of retention in verbal-learning situations. The influence of the kind of theoretical work done by Underwood (1957) and Underwood and Postman (1960) is beginning to be felt.

A variety of studies of the retention of motor skills were conducted during the third era. For the early part of the period, these studies may be separated into two varieties: those concerned with retention, as such, and those concerned with the warming-up phenomenon. Very recently, a third type of study that attempts to analyze the factors of memory has emerged.

A number of studies using the pursuit rotor were conducted. These employed retention intervals ranging from 1 week (Jahnke, 1958) through 4 weeks (Jahnke & Duncan, 1956) and 1 year (Bell, 1950; Eysenck, 1960), to 2 years (Koonce, Chambliss, & Irion, 1964). All studies showed high degrees of retention and rapid recovery of retention losses (if any) in relearning. Both the Jahnke and Duncan study and the one by Koonce *et al.* revealed that reminiscence effects might be observed after relatively long retention intervals (4 weeks and 2 years, respectively). Eysenck's study was noteworthy for the fact that he gave a great deal of prerest practice (fifty 15-minute sessions), which probably precluded obtaining reminiscence owing to the fact that his subjects were approaching perfect scores by the end of initial practice. He found an average retention loss of about 10% in mean time-on-target score.

In other, generally more complicated, learning situations, a high degree of retention has also been demonstrated, although habits involving discrete motor responses seem to be less well remembered than habits involving continuous responses (Adams & Hufford, 1962). Representative studies include those by E. I. Jones and Bilodeau (1953), Battig, Nagel, Voss, and Brogden (1957), and Ammons *et al.* (1958). Naylor and Briggs have reviewed the literature on long-term retention of learned skills (1961).

The warming-up effect was not discovered as a phenomenon or born as a concept during the third period of our history. The effect had been noted in very early times, particularly in connection with work phenomena, and had been specifically linked to practice situations by Thorndike (1914). In their study of the warming-up effect, Robinson and Heron (1924) defined it in the following terms: "Warming-up is a rise in efficiency which is steeper and more temporary than the rise which can be seen in, let us say, successive daily performances" (p.81). Interest in the warming-up phenomenon in relation to problems of memory was restimulated during the third period by the emphasis placed upon this concept by Ammons (1947a) and Irion (1948).[10]

Amount of warming-up as a function of the number of times the activity has been resumed can be studied in distribution-of-practice situations by comparing the postrest performance curves that follow successive periods of rest. If the results of Adams (1952) and Barch (1954) are considered together, it seems probable that the amount of the warming-up effect, as measured on such successive resumptions of practice, at first increases and then decreases. Warming-up as a function of length of interpolated rest can be studied in reminiscence experiments that vary length of rest. The results of Jahnke and Duncan (1956) and Koonce *et al.* (1964) suggest that, as length of rest increases, the magnitude of the warming-up effect at first increases and then decreases.

Attempts either to influence the level of initial performance or to reduce the amount of warm-up decrement present in postrest performance by introducing an independent warming-up activity have met with mixed success. Ammons (1951b) failed to obtain any indication of improved initial or postrest rotor performance as a result of using any one of several kinds of prepractice activities. Walker, DeSoto, and Shelly (1957) also report negative results. In this investigation, right-handed rotoring failed to serve as an effective warming-up activity for left-handed rotoring. On the other hand, Silver (1952) used the reversed alphabet task and found that postrest performance was facilitated by prerecall warming-up. He found that amount of facilitation increased and then decreased as a function of amount of massed warming-up activity, that the optimum amount of massed warming-up depended upon the amount of prerest practice, and that subjects could tolerate greater amounts of distributed warming-up than massed warming-up. These findings are generally more compatible with results obtained in verbal retention situations, but perhaps they result from the fact that the alphabet printing task more resembles verbal learning situations than it resembles the rotoring situation. Perhaps, and for reasons unknown, it is not possible to reinstate facilitative sets in learning situations that involve small amounts of response selectivity or high amounts of response unavailability, or both.

---

[10] Much of the early work on warming-up has been reviewed by Robinson (1934). The relationship between warming-up and retention has been reviewed in detail by Adams (1961b).

The latest trend in the study of motor memory is represented by four very recent papers (E. A. Bilodeau, Sulzer & Levy, 1962; E. A. Bilodeau, Levy, & Sulzer, 1963; E. A. Bilodeau & Levy, 1964; E. A. Bilodeau, Jones, & Levy, 1964). These studies employ a very simple (lever positioning) response, and extremely limited amounts of training (two guided responses plus one "free," or training, response) which may or may not be followed by immediate or delayed knowledge of results, which knowledge of results is independent (in magnitude) of the response made by the subject. Both the response that the subject makes (as well as the one he doesn't choose) and the knowledge of results given to him can be scaled in magnitude. The subject's ability to duplicate the response that he made on his "free" response trial, his ability to produce the response he did not select on that trial, and the accuracy of his memory for the knowledge of results that was given to him (if any) can be measured, and the relationships between these measures can be studied. In addition, the values of these measures as a function of length of retention interval, knowledge of results administered, delay of knowledge of results, number of previous measurements of these variables, etc. can be studied. While it is probably too early to evaluate the contributions that these studies make to our knowledge of retention in any detailed way, the methodological implications of this work are crystal clear. Not only are different components separately analyzable, but the use of measures other than the usual mean score represents a considerable advance. Forgetting may be shown not only by a falling mean score, but also by changes in variability and deterioration of the correlation of between-trial performances. It is perhaps not too much to hope that these studies will open the door to a new dimension of research on problems of retention.

The problem of knowledge of results or feedback has had a long and rather involved history, but the overwhelming majority of studies of feedback were conducted during the third period.

The early phase of the work on knowledge of results, at least as this concept applies to skill-learning situations, flows more or less directly from

Thorndike's (1911) statement of the law of effect. By the late 1930's, a considerable number of studies had been conducted which showed that knowledge of results was an effective variable: that little improvement occurred without knowledge of results, that the introduction of knowledge of results resulted in improvement, and that withdrawal of knowledge of results was followed by a deterioration of performance.[11]

Most of the early studies of knowledge of results were concerned principally with the effects of giving or withholding knowledge of results rather than with an analysis of the effects that could be produced by varying the nature of the administration of it. Also, the results of animal experimentation on reinforcement and the formulation of reinforcement theory by Hull (1943) exerted a strong influence on the general conception of the effects of knowledge of results. This is to say that there was a considerable tendency to accept the findings of animal research on problems of reinforcement as being applicable to the feedback problem in human verbal and motor learning. In this instance, at least, the influence of animal research and Hullian theory was unfortunate.

A return to the study of knowledge of results in the context of human learning was initiated by Brown (1949), who was working in the context of Hullian theory but whose analysis, in effect, achieved a *rapprochement* between the views of Hull and the earlier conceptions of Thorndike. Brown's paper, which was a proposal for a program of research on knowledge of results, exerted its influence principally through the researches of the Bilodeaus that are cited in the following paragraphs.

A number of problems have been investigated in the field of knowledge of results. These include (1) delay of knowledge of results, (2) the effects of filling the delay period in such experiments in various ways, (3) the frequency of knowledge of results, and (4) transformations of knowledge of results.[12]

---

[11] Representative studies from this period include Thorndike (1927), Lorge and Thorndike (1935), and Elwell and Grindley (1938). Thorndike (1931, 1932) summarized his theoretical position regarding knowledge of results and reviewed many of his experimental findings.

---

[12] General summaries of portions of the work in these areas will be found in Wolfle (1951), Ammons (1956), E. A. Bilodeau (1956b), E. A. Bilodeau and Bilodeau (1961), and Adams (1964). A related problem, not considered here, is the effect of distorted, diminished, or delayed feedback upon the performance of skillful acts that are relatively continuous and that have already been acquired. Aside from such anticipations as Aristotle's illusion, the first work in this area would appear to have been done by Stratton (1896, 1897). The work of Fairbanks (1955), Bahrick (1957), and Chase and his collaborators (1961a,b) deserves to be mentioned. The most extensive and impressive recent work on this problem has been undertaken by Smith and his associates (see Smith, 1961, 1962; Smith & Smith, 1962; Smith *et al.*, 1963, 1964).

In contrast to the findings concerning delay of reinforcement in animal studies, the general conclusion that can be reached concerning the effects of delay of knowledge of results is that, within very wide limits, delay of knowledge of results has little effect upon the learning of discrete tasks provided that the subject is not required to make other responses of the same class during the delay period. This conclusion is in general agreement with the classical finding of Lorge and Thorndike (1935). Intervals of delay of from .25 seconds (Noble & Alcock, 1958) through 1 week (Ryan & Bilodeau, 1962) to 4 weeks (E. A. Bilodeau *et al.*, 1962) have been used without obtaining impressive evidence for the degrading effects of delay (see also, E. A. Bilodeau & Bilodeau, 1958). It should be emphasized that these conclusions do not apply to continuous tracking tasks where delay or lag in informative feedback can have seriously disruptive effects upon learning and performance (Conklin, 1957, 1959; Garvey, Sweeney, & Birmingham, 1958). In the population of discrete tasks, the experimental analog to the situation that applies in the case of continuous tasks is the situation in which knowledge of results is delayed for a specific number of stimulus-response sequences rather than for a particular length of (unfilled) time. In such situations, the subject is given knowledge of results for his performance on some specified, but not immediately preceding, trial. This is the situation that was used by Lorge and Thorndike (1935) and more recently by I. McD. Bilodeau (1956). Under these circumstances, delay is disruptive of performance and learning.

In a study of learning as a function of frequency of knowledge of results, E. A. Bilodeau and Bilodeau (1958) compared the learning of groups that received 100, 33, 25, and 10 percent knowledge of results in a lever positioning task. Although, as the authors point out, there are numerous differences between the partial reinforcement situation as it has been studied in animal experiments and the "partial knowledge of results" procedures used in their experiment, their finding that "learning is independent of relative frequency and positively related to absolute frequency" (of KR, knowledge of results) does not seem to mesh very well with the results of the animal experiments as these are usually interpreted. More recently, Taylor and Noble (1962) have obtained somewhat similar results using 100, 75, 50 and 25 percent knowledge of results in the Selective Mathometer situation. They found that their 100-percent group was best in acquisition (when plotted against either number of trials or number of reinforcements). There was some suggestion that the 100-percent group was slightly more susceptible to extinction, although this result was confined to the early extinction trials. More work needs to be done on this problem.

There have been a considerable number of studies involving transformation of knowledge of results. The simplest of these transformations involves varying target size in pursuit or tracking situations. E. A. Bilodeau (1956b) has discussed the essential aspects of this problem. Over a considerable range of target sizes, target size does not appear to affect learning. It should be noted that target size can be increased as well by transforming the subject's true score into a reported feedback that contains a good bit of latitude for response error as by increasing the physical size of the target button on a pursuit rotor. The latter procedure has the advantage that the recorded score and the reported score are not both affected by the "size" of the target—a factor that seems to account for all of the differences in performance of subjects using different sizes of physical targets. Using such transformations, it can be shown that drastic reductions or increases in the error reported as feedback can have a deleterious effect upon performance, with reduction having more effect than increase (if the subject is *always* informed that he "hit the target," his performance fails to improve). For example, using the Manual Lever Test, E. A. Bilodeau (1954b) transformed true scores by making them 0, 2, 4, 8, or 16 units closer to the goal score than they actually were (in separate groups). In this situation, the 0 group was given its true score as feedback, and the 2, 4, 8, and 16 groups worked with successively broader targets. The results indicated that the groups with the narrower targets learned more rapidly than the groups with the broader targets. In a second study, E. A. Bilodeau (1955b) compared transformations that reduced error by various amounts to transformations that increased error by the same amounts. Again, reduced error transformations resulted in poorer performance with the greater degrees of transformation. Transformations that reduced or increased error were approximately equal in their effects with small transformations, but with the larger transformations, reduced error resulted in a considerably poorer performance than increased error.[13]

Linear transformations of knowledge of results have also been studied (E. A. Bilodeau, 1953), the general finding being that transformations that increasingly minimize error do not affect final performance but do, increasingly, degrade the rate of approach to the common asymptote. On the other hand, Noble and Broussard (1955) have shown that knowledge of results transformed to inflate reported error has the effect of making the subject "overshoot" the target. This is followed by a rapid reapproach to zero error, but with some evidences of "hunting." In this case, the variability of response early in practice is much larger for the transformation that inflates error than for the transformation that deflates error.

---

[13] Such results can be expected only when the subject is given feedback that specifies degree and/or direction of error. If the subject is only told "hit" or "miss," it would be expected that a U-shaped function would be obtained when amount of learning is plotted against target width. With extremely small targets (of the invisible variety we are discussing), the subject would spend an inordinate amount of practice time locating them. With extremely broad targets, so much information would be lost in the feedback that it would cease to be "informative" feedback. Results that suggest such a U-shaped function have been obtained by Battig, Nagel, and Brogden (1955), but in a bidimensional compensatory tracking situation that is so complicated as to make interpretation difficult.

It is evident that, in the case of many skill-learning tasks, knowledge of results is intrinsic to the task to a greater or lesser degree. In learning to shoot skeet, for example, or in learning to play the violin, the results of responding are fairly evident to the eye or to the ear. The clarity of the knowledge of results that is intrinsic to the learning task may be the variable that underlies some of the contradictory findings in the case of augmented feedback experiments. Where clarity is low, augmented feedback may become a "crutch" that benefits performance only, with the result that performance deteriorates to the level of nonaugmented training when the additional knowledge of results is withdrawn. In situations where intrinsic clarity is high, augmentation of knowledge of results may tend to benefit both learning and immediate performance. This, at least, would seem to be the conclusion that Adams (1964) draws from results obtained by Kinkade (1963), a conclusion that he uses to resolve the contradictory results obtained in a number of different training situations.[14] Not to get drawn off into the argument concerning the use of red filters in the pedestal sight as a source of augmented feedback, it would still seem that the matter is not entirely settled owing to the unfortunate disparity of results obtained in the pursuit-rotor situation. Briefly, Reynolds and Adams (1953) obtained permanent effects from using augmented feedback ("click reinforcement") while neither Archer and Namikas (1958) nor I. McD. Bilodeau and Rosenquist (1964) were able to demonstrate the beneficial effects of augmented feedback. While various hypotheses might be advanced to account for these differences in results (composition of the subject populations, use of a tone rather than a click for augmented feedback purposes, etc.), to accept any of them would seriously limit the generality of the basic principle. If the Reynolds and Adams results are discounted, there is no need for the hypothesis since it was devised mainly for the purpose of explaining the different effects of augmented feedback in the "clear" rotor situation and in the "unclear" pedestal sight situation. However, the only possible basis for ignoring the findings of Reynolds and Adams would be to assume that they were contaminated by errors of random sampling, a contamination that is bound to occur with predictable frequency, but a source of error that is not suggested by the internal consistency of their data. For those who are unwilling to adopt the sampling-error assumption, a problem remains to be solved.

---

[14] In their thoughtful analysis, Annett and Kay (1957) came to somewhat different conclusions regarding the problem of augmented feedback. Using Miller's (1953) distinction between action feedback and learning feedback, they concluded that augmented feedback should tend to benefit performance (but not learning) in the case of action feedback, but should tend to affect both learning and performance in the case of learning feedback.

Whatever the confusions and misinterpretations may be in the data or in the theory, there can be small doubt that knowledge of results is the single most important variable governing the acquisition of skillful habits.

As far as the psychometric approach to the problem of skills is concerned, the third period started off with the workers in this area possessing a number of advantages over the workers of earlier periods. For one thing, an enormous amount of effort had been devoted during World War II to the production of tests that measured psychomotor abilities or, at least, were predictive of various psychomotor performances. Although many of these tests had been developed as a result of a considerable amount of trial-and-error behavior on the part of the psychometric specialists, a goodly number of tests were developed and the problem-solving type of behavior gave rise to new techniques for solving problems in the psychometric field.

If we put aside the great number of more or less routine standardization and validation studies, one of the earliest problems to be attacked during the third period was the relative power of tests to predict psychomotor performance early and late in practice. For example, Reynolds (1952) correlated a considerable number of test scores with performance on the Complex Coordinator at various stages of practice on the latter instrument. He found that the predictive value of the tests decreased as practice on the Complex Coordinator increased. This finding might suggest that later performance could better be predicted from early performance on the same task than by the use of "outside" tests were it not for the usual finding that initial scores become less and less predictive of final scores as amount of practice increases. Reynolds suggested that, as practice increased, the abilities demanded in performance might shift to those that were more specific to the particular task at hand. However, this hypothesis was not given much support by the findings of Adams (1953, 1957) or by the later findings of Parker and Fleishman (1961) which showed that selected outside tests, presumably those that measure abilities relevant to performance later in practice, were more predictive of final performances than were scores taken from the same task early in practice. In commenting on Adams' (1953) findings, Fleishman and Hempel (1955) remarked:

> "Moreover, the best predictors of advanced performance were measures of advanced performances on other psychomotor tasks. This would seem to indicate the presence of common factors at advanced stages of practice on

several tasks not held in common between these tasks earlier in practice." (p. 311).

However, in fairness to the specificity hypothesis of Reynolds, it should be noted that Fleishman and Hempel (1954, 1955) found that the loadings of some task-specific factors (in the Complex Coordinator and the Discrimination Reaction Time tests) increase with practice.

From such findings, it would seem to follow that practice might be regarded as resulting in a change in the pattern of abilities being employed in performance. Such an interpretation was made by Fleishman (1956) and by Fleishman and Hempel (1955). Whether the change in pattern is due to a relatively simple change in factor structure as suggested by Woodrow (1938b), a gain in complication of structure as suggested by the results obtained by Fleishman and Hempel (1955) or to a simplification as urged by M. B. Jones (1962), would seem to be a matter that awaits the accumulation of greater amounts of relevant data. Certainly, the simplicial model seems to be most congruent with the kinds of anecdotes that workers in the skills field have been telling each other about motor-learning behavior for many years. On the other hand, it might be possible that the basic nature of the change of factor structure that occurs with practice is dependent upon the nature of the learning task—a possibility that is not particularly pleasant to contemplate.

A series of studies related to this problem should be mentioned in passing. These investigations concern the fate under practice of groups that have been fractionated according to initial ability. Following the logic imposed by the decreasing correlation between trial one and subsequent trials (and the notion that the patterning of a set of independently variable abilities is systematically changing as a function of practice), it has been assumed that such groups will converge as practice progresses. However, if $x$ represents an individual's score early in practice and $y$ represents his score late in practice (both $x$ and $y$ being expressed in deviation form), then, according to a common form of the regression equation, $y = r\,(\sigma_y/\sigma_x)\,x$. If the value of $r$ exhibits a systematic decrease, and if the ratio of $\sigma_y/\sigma_x$ remains constant, then $y$ should, indeed, be less deviant than $x$ and the scores should converge. Although usually variability shows a tendency to decrease with practice in motor learning (which would enhance the convergence effect), this need not always be the case. Should variability increase with practice, the ratio of $\sigma_y/\sigma_x$ will become increasingly larger than 1.00, which could compensate (or more than compensate) for the dropping value of $r$, thereby producing curves that were parallel or divergent.

As far as the results are concerned, Adams (1957) found convergence but Reynolds and Adams (1954) obtained parallel curves under both massed and distributed practice conditions while Zeaman and Kaufman (1955) obtained converging curves under massed trials and parallel curves under distributed trials.[15] Although the data from the Reynolds and Adams (1954) study are not presented in such a form as to make a trial-by-trial comparison of variability possible (and although the evidence that is presented would not suggest a systematic increase in variability), it would appear to be almost necessary to assume such an increase in order to account for their parallel acquisition curves in the face of a falling coefficient of correlation. This is evidently what happened in the case of the distributed practice condition of the Zeaman and Kaufman (1955) experiment. Their Figure 2 reveals that the fractionated groups improved in a parallel fashion, but their Figure 7 shows that, as the correlation between first trial score and later scores dropped, variability of the group increased. More work needs to be done on this problem with particular care being devoted to the method of fractionation and with means, standard deviations, and correlations being reported for the same data.

In spite of the somewhat indeterminate results that have been obtained from ordinary fractionation experiments, the basic concept is considerably strengthened by the results obtained by Fleishman and Hempel (1955) showing that groups fractionated on the basis of different predictor variables converge or diverge under practice depending on the importance of the predictor variables early and late in practice. Despite the methodological criticisms that have been made by Bechtoldt (1962), M. B. Jones (1962), and Adams (1964), the attractiveness of the basic conception remains. Fleishman and Hempel (1955) expressed this view in the following terms: "Performance at any stage of practice is determined by a set of cooperating but independently variable abilities. Continuation of practice on the task is regarded as a change in the conditions under which these various abilities operate" (p. 311). Their interpretation is, of course, very similar to the conclusion reached by Woodrow (1938a) nearly 20 years earlier. If the intervening years have not advanced understanding greatly, they have given a much greater degree of empirical support to this kind of interpretation. Perhaps with the emergence of new approaches to the problem of correlation analysis (Guttman, 1954; M. B. Jones, 1959, 1960, 1962) new dimensions of understanding will also evolve.

---

[15] Reynolds and Adams (1954) interpret their parallel curves as being consistent with Hull's (1943, 1945) assumption that individual difference variables affect only the constants of a learning function. Actually as Zeaman and Kaufman (1955) note, a variation in Hull's (1943) rate-of-learning constant (*i*) leads to the prediction that the curves should diverge and then, with continued practice, come together again (since $M$ is independent of *i*). Zeaman and Kaufman interpret their parallel learning curves for their distributed practice condition by assuming that individual differences result from different levels of pre-experimentally produced conditioned inhibition, $_sI_R$.

A few closing remarks should be offered. The study of history does not appear to have had any very significant effect upon the work of the active investigators in the field of motor skills. One does not gain the impression that very many research workers have spent long hours delving into the musty archives that contain the records of the past. From the recency of the references cited in the experimental literature, and from the generally linear progress that seems to have occurred in the various investigative fields, one draws the conclusion that a particular research finding is either assimilated into the body of working knowledge almost immediately, or is discarded with relative finality. While it is true that historical research seems to have flourished more in the past than at the present time, and while an occasional investigator sometimes pauses even now to look over his shoulder at the past, is also fairly apparent that most contemporary experimenters are precariously balanced on the knife edge of the present and leaning toward the future and that historical treatments, such as the one offered here, are essentially anachronistic.

## REFERENCES

Adams, J. A. Warm-up decrement in performance on the pursuit-rotor. *Amer. J. Psychol.*, 1952, **65**, 404–414.

Adams, J. A. The prediction of performance at advanced stages of training on a complex psychomotor task. *USAF Hum. Resour. Res. Cent. res. Bull.*, 1953, No. 53–49.

Adams, J. A. Psychomotor performance as a function of intertrial rest interval. *J. exp. Psychol.*, 1954, **48**, 131–133.

Adams, J. A. A source of decrement in psychomotor performance. *J. exp. Psychol.*, 1955, **49**, 390–394.

Adams, J. A. The relationship between certain measures of ability and the acquisition of a psychomotor criterion response. *J. gen. Psychol.*, 1957, **56**, 121–134.

Adams, J. A. Human tracking behavior. *Psychol. Bull.*, 1961, **58**, 55–79. (a)

Adams, J. A. The second facet of forgetting: a review of warm-up decrement *Psychol. Bull.*, 1961, **58**, 257–273. (b)

Adams, J. A. Test of the hypothesis of psychological refractory period. *J. exp. Psychol.*, 1962, **64**, 280–287.

Adams, J. A. Motor skills. *Annu. Rev. Psychol.*, 1964, **15**, 181–202.

Adams, J. A., & Hufford, L. E. Contributions of a part-task trainer to the learning and relearning of a time-shared flight maneuver. *Hum. Factors*, 1962, **4**, 159–170.

Adams, J. A., & Reynolds, B. Effect of shift in distribution of practice conditions following interpolated rest. *J. exp. Psychol.*, 1954, **47**, 32–36.

Ammons, R. B. Acquisition of motor skill: I. Quantitative analysis and theoretical formulation. *Psychol. Rev.*, 1947, **54**, 263–281. (a)

Ammons, R. B. Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–411. (b)

Ammons, R. B. Effect of distribution of practice on rotary pursuit 'hits.' *J. exp. Psychol.*, 1951, **41**, 17–22. (a)

Ammons, R. B. Effects of pre-practice activities on rotary pursuit performance. *J. exp. Psychol.*, 1951, **41**, 187–191. (b)

Ammons, R. B. Effects of knowledge of performance: a survey and tentative theoretical formulation. *J. gen. Psychol.*, 1956, **54**, 279–299.

Ammons, R. B., & Willig, L. Acquisition of motor skill: IV. Effects of repeated periods of massed practice. *J. exp. Psychol.*, 1956, **51**, 118–126.

Ammons, R. B., Ammons, C. H., & Morgan, R. L. Transfer of skill and decremental factors along the speed dimension in rotary pursuit. *Percept. mot. Skills*, 1956, **6**, 43.

Ammons, R. B., Farr, R. G., Bloch, E., Neumann, E., Dey, M., Marion, R. & Ammons, C. H. Long-term retention of perceptual-motor skills. *J. exp. Psychol.*, 1958, **55**, 318–328.

Andreas, B. G., Green, R. F., & Spragg, S. D. S. Transfer effects between performance on a following tracking task (Modified SAM Two-Hand Coordination Test) and a compensatory tracking task (Modified SAM Two-Hand Pursuit Test). *J. Psychol.*, 1954, **37**, 173–183.

Annett, J., & Kay, H. 'Skilled performance.' *Occup. Psychol.*, 1956, **30**, 112–117.

Annett, J., & Kay, H. Knowledge of results and 'skilled performance.' *Occup. Psychol.*, 1957, **31**, 69–79.

Archer, E. J. Effect of distribution of practice on a component skill of rotary pursuit tracking. *J. exp. Psychol.*, 1958, **56**, 427–436.

Archer, E. J., & Namikas, G. A. Pursuit rotor performance as a function of delay of information feedback. *J. exp. Psychol.*, 1958, **56**, 325–327.

Attneave, F. *Applications of information theory to psychology*. New York: Holt, 1959.

Bahrick, H. P. An analysis of stimulus variables influencing the proprioceptive control of movements. *Psychol. Rev.*, 1957, **64**, 324–328.

Bahrick, H. P., Fitts, P. M., & Briggs, G. E. Learning curves – facts or artifacts? *Psychol. Bull.*, 1957, **54**, 256–268.

Baker, K. E., & Wylie, R. C. Transfer of verbal training to a motor task. *J. exp. Psychol.*, 1950, **40**, 632–638.

Baker, K. E., Wylie, R. C., & Gagné, R. M. Transfer of training to a motor-skill as a function of variation in rate of response. *J. exp. Psychol.*, 1950, **40**, 721–732.

Barch, A. M. The effect of difficulty of task on proactive facilitation and interference. *J. exp. Psychol.*, 1953, **46**, 37–42.

Barch, A. M. Warm-up in massed and distributed pursuit rotor performance. *J. exp. Psychol.*, 1954, **47**, 357–361.

Barch, A. M., & Lewis, D. The effect of task difficulty and amount of practice on proactive transfer. *J. exp. Psychol.*, 1954, **48**, 134–142.

Battig, W. F., Nagel, E. H., & Brogden, W. J. The effects of error-magnification and marker-size on bidimensional compensatory tracking. *Amer. J. Psychol.*, 1955, **68**, 585–594.

Battig, W. F., Nagel, E. H., Voss, J. F., & Brogden, W. J. Transfer and retention of bidimensional compensatory tracking after extended practice. *Amer. J. Psychol.*, 1957, **70**, 75–80.

Bechtoldt, H. P. Factor analysis and the investigation of hypotheses. *Percept. mot. Skills*, 1962, **14**, 319–342.

Bell, H. M. Rest pauses in motor learning as related to Snoddy's hypothesis of mental growth. *Psychol. Monogr.*, 1942, **54**, No. 1 (Whole No. 243).

Bell, H. M. Retention of pursuit rotor skill after one year. *J. exp. Psychol.*, 1950, **40**, 648–649.

Bilodeau, E. A. Decrements and recovery from decrements in a simple work task with variation in force requirements at different stages of practice. *J. exp. Psychol.*, 1952, **44**, 96–100. (a)

Bilodeau, E. A. Massing and spacing phenomena as functions of prolonged and extended practice. *J. exp. Psychol.*, 1952, **44**, 108–113. (b)

Bilodeau, E. A. Speed of acquiring a simple motor response as a function of the systematic transformation of knowledge of results. *Amer. J. Psychol.*, 1953, **66**, 409–420.

Bilodeau, E. A. Rate recovery in a repetitive motor task as a function of successive rest periods. *J. exp. Psychol.*, 1954, **48**, 197–203 (a)

Bilodeau, E. A. Accuracy of response as a function of target width. *J. exp. Psychol.*, 1954, **47**, 201–207. (b)

Bilodeau, E. A. Variations in knowledge of component performance and its effects upon part-part and part-whole relations. *J. exp. Psychol.*, 1955, **50**, 215–224. (a)

Bilodeau, E. A. Motor performance as affected by magnitude and direction of error contained in knowledge of results. *J. Psychol.*, 1955, **40**, 103–113. (b)

Bilodeau, E. A. Transfer of training after part practice on a dual-control tracking apparatus. *USAF Personn. Train. Res. Cent. res. Rep.*, 1956, No. 56–110. (a)

Bilodeau, E. A. Studies of target size and the control of psychomotor behavior through systematic transformation of knowledge of results. In, G. Finch & R. Cameron (Eds.), *Symposium on Air Force human engineering, personnel, and training research*, Baltimore, Md.: Air Research and Development Command, 1956, (ARDC Tech. Rep. No. 56–8). (b)

Bilodeau, E. A. The relationship between a relatively complex motor skill and its components. *Amer J. Psychol.*, 1957, **70**, 49–55. (a)

Bilodeau, E. A. Patterns of internal consistency in multipart skilled performances. *Amer. J. Psychol.*, 1957, **70**, 550–559. (b)

Bilodeau, E. A., & Bilodeau, I. McD. The contribution of component activities to the total psychomotor task. *J. exp. Psychol.*, 1954, **47**, 37–46.

Bilodeau, E. A., & Bilodeau, I. McD. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612.

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Bilodeau, E. A., & Levy, C. M. Long-term memory as a function of retention time and other conditions of training and recall. *Psychol. Rev.*, 1964, **71**, 27–41.

Bilodeau, E. A., & Ryan, F. J. Prediction of complex task proficiency by means of component responses. *Percept. mot. Skills*, 1961, **12**, 299–306.

Bilodeau, E. A., Sulzer, J. L., & Levy, C. M. Theory and data on the interrelationships of three factors of memory. *Psychol. Monogr.*, 1962, **76**, No. 20 (Whole No. 539).

Bilodeau, E. A., Levy, C. M., & Sulzer, J. L. Long-term retention under conditions of artificially induced recall of related events. *Percept. mot. Skills*, 1963, **16**, 895–910; *Monogr. Suppl.* No. 8-V16.

Bilodeau, E. A., Jones, M. B., & Levy, C. M. Long-term memory as a function of retention time and repeated recalling. *J. exp. Psychol.*, 1964, **67**, 303–309.

Bilodeau, I. McD. Accuracy of a simple positioning response with variation in the number of trials by which knowledge of results is delayed. *Amer. J. Psychol.*, 1956, **69**, 434–437.

Bilodeau, I. McD., & Rosenquist, H. S. Supplementary feedback in rotary-pursuit tracking. *J. exp. Psychol.*, 1964, **68**, 53–57.

Bourne, L. E., Jr., & Archer, E. J. Time continuously on target as a function of distribution of practice. *J. exp. Psychol.*, 1956, **51**, 25–33.

Bray, C. W. Transfer of learning. *J. exp. Psychol.*, 1928, **11**, 443–467.

Britt, S. H. Retroactive inhibition: A review of the literature. *Psychol. Bull.*, 1935, **32**, 381–440.

Brown, J. S. A proposed program of research on psychological feedback (knowledge of results) in the performance of psychomotor tasks. In, Research planning conference on perceptual and motor skills, *USAF Hum. Resour. Res. Center Conf. Rep.* 31, 1949, No. 49–2. Pp. 81–87.

Bruce, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343–361.

Bryan, W. L., & Harter, N. Studies in the physiology and psychology of the telegraphic language. *Psychol. Rev.*, 1897, 4, 27–53.

Bryan, W. L., & Harter, N. Studies on the telegraphic language. The acquisition of a hierarchy of habits. *Psychol. Rev.*, 1899, **6**, 345–375.

Buxton, C. E. The application of multiple factorial methods to the study of motor abilities. *Psychometrika*, 1938, **3**, 85–93.

Buxton, C. E. Retroaction and gains in motor learning, III: evaluation of results. *J. gen. Psychol.*, 1940, **22**, 309–320.

Buxton, C. E. Reminiscence in the acquisition of skill. *Psychol. Rev.*, 1942, **49**, 191–196.

Buxton, C. E. The status of research in reminiscence. *Psychol. Bull.*, 1943, **40**, 313–340.

Buxton, C. E., & Grant, D. A. Retroaction and gains in motor learning, II: sex differences, and a further analysis of gains. *J. exp. Psychol.*, 1939, **25**, 198–208.

Buxton, C. E., & Henry, C. E. Retroaction and gains in motor learning, I: similarity of interpolated task as a factor in gains. *J. exp. Psychol.*, 1939, **25**, 1–17.

Buxton, C. E., & Humphreys, L. G. The effect of practice upon intercorrelations of motor skill. *Science*, 1935, **81**, 441–442.

Campbell, M. The 'personal equation' in serial pursuit performances. *J. appl. Psychol.*, 1934, **18**, 785–792.

Chase, R. A., Harvey, S., Standfast, S., Rapin, I., & Sutton, S. Studies on sensory feedback, I: effect of delayed auditory feedback on speech and keytapping. *Quart. J. exp. Psychol.*, 1961, **13**, 141–152. (a)

Chase, R. A., Rapin, I., Gilden, L., Sutton, S., & Guilfoyle, G. Studies on sensory feedback, II: Sensory feedback influences on keytapping motor tasks. *Quart. J. exp. Psychol.*, 1961, **13**, 153–167. (b)

Conklin, J. E. Effect of control lag on performance in a tracking task. *J. exp. Psychol.*, 1957, **53**, 261–268.

Conklin, J. E. Linearity of the tracking performance function. *Percept. mot. Skills*, 1959, **9**, 387–391.

Courts, F. Relations between muscular tension and performance. *Psychol. Bull.*, 1942, **39**, 347–368.

Crafts, L. W. Whole and part methods with unrelated reactions. *Amer. J. Psychol.*, 1930, **42**, 591–601.

Crafts, L. W. Transfer as related to number of common elements. *J. gen. Psychol.*, 1935, **13**, 147–158.

Day, R. H. Relative task difficulty and transfer of training in skilled performance. *Psychol. Bull.*, 1956, **53**, 160–168.

Digman, J. M. Performance under optimal practice conditions following three degrees of massing of early practice. *J. exp. Psychol.*, 1956, **52**, 189–193.

Digman, J. M. Growth of a motor skill as a function of distribution of practice. *J. exp. Psychol.*, 1959, **57**, 310–316.

Doré, L. R., & Hilgard, E. R. Spaced practice and the maturation hypothesis. *J. Psychol.*, 1937, **4**, 245–259.

Doré, L. R., & Hilgard, E. R. Spaced practice as a test of Snoddy's two processes in mental growth. *J. exp. Psychol.*, 1938, **23**, 359–374.

Duncan, C. P. Transfer in motor learning as a function of degree of first-task learning and inter-task similarity. *J. exp. Psychol.*, 1953, **45**, 1–11.

Duncan, C. P. On the similarity between reactive inhibition and neural satiation. *Amer. J. Psychol.*, 1956, **69**, 227–235.

Duncan, C. P. Visual and kinesthetic components of reactive inhibition. *Amer. J. Psychol.*, 1957, **70**, 616–619.

Ellson, D. G. The independence of tracking in two and three dimensions with the G.E. Pedestal Sight. In G. Finch (Ed.), *Memorandum rep.* No. TSEAA-694-2G. AAF Air Materiel Command, Eng. Div., 1947.

Elwell, J. L., & Grindley, G. C. The effect of knowledge of results on learning and performance: I. A co-ordinated movement of the two hands. *Brit. J. Psychol.*, 1938, **29**, 39–54.

Espenschade, A. Motor performance in adolescence. *Monogr. Soc. Res. Child Develpm.*, 1940, **5**, No. 1.

Eysenck, S. B. J. Retention of a well-developed motor skill after one year. *J. gen. Psychol.*, 1960, **63**, 267–273.

Fairbanks, G. Selective vocal effects of delayed auditory feedback. *J. Speech & Hear. Disord.*, 1955, **20**, 333–345.

Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.

Fitts, P. M. Engineering psychology. *Annu. Rev. Psychol.*, 1958, **9**, 267–294.

Fleishman, E. A. Predicting advanced levels of proficiency in psychomotor skills. In G. Finch & R. Cameron (Eds.), *Symposium on Air Force human engineering, personnel, and training research*, Baltimore, Md.: Air Research and Development Command, 1956, (ARDC Tech. Rpt. No. 56–8).

Fleishman, E. A., & Hempel, W. E., Jr. Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika*, 1954, **19**, 239–252.

Fleishman, E. A., & Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. *J. exp. Psychol.*, 1955, **49**, 301–312.

Freeman, F. N., & Abernethy, E. M. Comparative retention of typewriting and of substitution with analogous material. *J. educ. Psychol.*, 1930, **21**, 639–647.

Freeman, F. N., & Abernethy, E. M. New evidence of the superior retention of typewriting to that of substitution. *J. educ. Psychol.*, 1932, **23**, 331–334.

Gagné, R. M., & Foster, H. Transfer of training from practice on components in a motor skill. *J. exp. Psychol.*, 1949, **39**, 47–68. (a)

Gagné, R. M., & Foster, H. Transfer to a motor skill from practice on a pictured representation. *J. exp. Psychol.*, 1949, **39**, 342–354. (b)

Gagné, R. M., Foster, H., & Crowley, M. E. The measurement of transfer of training. *Psychol. Bull.*, 1948, **45**, 97–130.

Gagné, R. M., Baker, K. E., & Foster, H. Transfer of discrimination training to a motor task. *J. exp. Psychol.*, 1950, **40**, 314–328.

Garvey, W. D., Sweeney, J. S., & Birmingham, H. P. Differential effects of 'display lags' and 'control lags' on the performance of manual tracking systems. *J. exp. Psychol.*, 1958, **56**, 8–10.

Gibbs, C. B. Transfer of training and skill assumptions in tracking tasks. *Quart. J. exp. Psychol.*, 1951, 3, 99–110.

Goldman, H. M. The loci of reminiscence in pursuit rotor learning as inferred from the nature of the interpolated rest activity. Unpublished master's thesis, Univer. of Illinois, 1952.

Goodenough, F. L., & Smart, R. C. Inter-relationships of motor abilities in young children. *Child Develpm.*, 1935, **6**, 141–153.

Grice, G. R., & Reynolds, B. Effect of varying amounts of rest on conventional and bilateral transfer 'reminiscence.' *J. exp. Psychol.*, 1952, 44, 247–252.

Guttman, L. A new approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences.* Glencoe, Ill.: Free Press, 1954.

Hicks, J. A. The acquisition of motor skill in young children: An experimental study of the effects of practice in throwing at a moving target. *Univ. Iowa Stud. Child. Welf.*, 1931, 4, No. 5.

Hilgard, E. R., & Smith, M. B. Distributed practice in motor learning: score changes within and between daily sessions. *J. exp. Psychol.*, 1942, **30**, 136–146.

Hovland, C. I., & Sears, R. R. Experiments on motor conflict: I. Types of conflict and their modes of resolution. *J. exp. Psychol.*, 1938, **23**, 477–493.

Hull, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.

Hull, C. L. The place of innate individual and species differences in a natural-science theory of behavior. *Psychol. Rev.*, 1945, **52**, 55–60.

Humphreys, L. G. The factor of time in pursuit rotor learning. *J. Psychol.*, 1937, 3, 429–436.

Irion, A. L. The relation of 'set' to retention. *Psychol. Rev.*, 1948, **55**, 336–341.

Irion, A. L. Reminiscence in pursuit-rotor learning as a function of length of rest and of amount of pre-rest practice. *J. exp. Psychol.*, 1949, **39**, 492–499.

Irion, A. L., & Gustafson, L. M. 'Reminiscence' in bilateral transfer. *J. exp. Psychol.*, 1952, **43**, 321–323.

Jahnke, J. C. Retention in motor learning as a function of amount of practice and rest. *J. exp. Psychol.*, 1958, **55**, 270–273.

Jahnke, J. C., & Duncan, C. P. Reminiscence and forgetting in motor learning after extended rest intervals. *J. exp. Psychol.*, 1956, **52**, 273–282.

Jones, E. I., & Bilodeau, E. A. Differential transfer of training between motor tasks of different difficulty. *USAF Hum. Resour. Res. Cent. res. Bull.*, 1952, No. 52–35.

Jones, E. I., & Bilodeau, E. A. Retention and relearning of a complex perceptual-motor skill after 10 months of no practice. *USAF Hum. Resour. Res. Cent. res. Bull.*, 1953, No. 53–17.

Jones, M. B. Simplex theory. *USN Sch. Aviat. Med. Monogr.*, 1959, No. 3.

Jones, M. B. Molar correlational analysis. *USN Sch. Aviat. Med. Monogr.*, 1960, No. 4.

Jones, M. B. Practice as a process of simplification. *Psychol. Rev.*, 1962, **69**, 274–294.

Jost, A. Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen. Z. *Psychol.*, 1897, 14, 436–472.

Kientzle, M. J. Properties of learning curves under varied distributions of practice. *J. exp. Psychol.*, 1946, **36**, 187–211.

Kientzle, M. J. Ability patterns under distributed practice. *J. exp. Psychol.*, 1949, **39**, 532–537.

Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.

Kimble, G. A. Transfer of work inhibition in motor learning. *J. exp. Psychol.*, 1952, **43**, 391–392.

Kimble, G. A., & Horenstein, B. R. Reminiscence in motor learning as a function of length of interpolated rest. *J. exp. Psychol.*, 1948, **38**, 239–244.

Kinkade, R. G. A differential influence of augmented feedback on learning and on performance. *USAF WADC tech. Rep.*, 1963, No. 63–12.

Köhler, W., & Fishback, J. The destruction of the Müller-Lyer illusion in repeated trials: II. Satiation patterns and memory traces. *J. exp. Psychol.*, 1950, **40**, 398–410.

Koerth, W. A pursuit apparatus: eye-hand coördination. *Psychol. Monogr.*, 1922, **31**, No. 1 (Whole No. 140).

Koonce, J. M., Chambliss, D. J., & Irion, A. L. Supplementary report: Long-term reminiscence in the pursuit-rotor habit. *J. exp. Psychol.*, 1964, **67**, 498–500.

Leavitt, H. J., & Schlosberg, H. The retention of verbal and of motor skills. *J. exp. Psychol.*, 1944, **34**, 404–417.

Lewis, D., McAllister, D. E., & Adams, J. A. Facilitation and interference in performance on the Modified Mashburn Apparatus: I. The effects of varying the amount of original learning. *J. exp. Psychol.*, 1951, **41**, 247–260.

Lewis, D., Smith, P. N., & McAllister, D. E. Retroactive facilitation and interference in performance on the Modified Two-Hand Coordinator. *J. exp. Psychol.*, 1952, **44**, 44–50.

Lewis, D., McAllister, D. E., & Bechtoldt, H. P. Correlational analysis of the learning and relearning of four different tasks on the Modified Mashburn Apparatus. *J. Psychol.*, 1953, **36**, 83–109. (a)

Lewis, D., McAllister, D. E., & Bechtoldt, H. P. Correlational study of performance during successive phases of practice on the standard and reversed tasks on the SAM Complex Coordinator. *J. Psychol.*, 1953, **36**, 111–126. (b)

Lincoln, R. S., & Smith, K. U. Transfer of training in tracking performance at different target speeds. *J. appl. Psychol.*, 1951, **35**, 358–362.

Lorge, I. Influence of regularly interpolated time intervals upon subsequent learning. *Teach. Coll. Contr. Educ.*, 1930, No. 438.

Lorge, I., & Thorndike, E. L. The influence of delay in the aftereffect of a connection. *J. exp. Psychol.*, 1935, **18**, 186–194.

Lyon, D. O. The relation of length of material to time taken for learning and the optimum distribution of time. *J. educ. Psychol.*, 1914, **5**, 1–9, 85–91, and 155–163.

McAllister, D. E., & Lewis, D. Facilitation and interference in performance on the Modified Mashburn Apparatus: II. The effects of varying the amount of interpolated learning. *J. exp. Psychol.*, 1951, **41**, 356–363.

McCloy, C. H. The influence of chronological age on motor performance. *Res. Quart. Amer. Ass. Hlth.*, 1935, **6**, 61–64.

McGeoch, J. A. The acquisition of skill. *Psychol. Bull.*, 1927, **24**, 437–466.

McGeoch, J. A. The comparative retention values of a maze habit, of nonsense syllables, and of rational learning. *J. exp. Psychol.*, 1932, **15**, 662–680.

McGeoch, J. A., & Irion, A. L. *The psychology of human learning* (2nd. ed.). New York: Longmans, Green, 1952.

McGeoch, J. A., & Melton, A. W. The comparative retention values of maze habits and of nonsense syllables. *J. exp. Psychol.*, 1929, **12**, 392–414.

McGinnis, E. The acquisition and interference of motor habits in young children. *Genet. Psychol. Monogr.*, 1929, **6**, 203–311.

Melton, A. W., (Ed.) *Apparatus tests.* Washington, D.C.: U.S. Gov. Printing Office, 1947. (AAF Aviat. Psychol. Prog. Res. Rep. No. 4). (a)

Melton, A. W., (Ed.) *Apparatus tests (supplement).* Washington, D.C.: U.S. Gov. Printing Office, 1947. (AAF Aviat. Psychol. Prog. Res. Rep. No. 4, Suppl.). (b)

Melton, A. W., & Briggs, G. E. Engineering psychology. *Annu. Rev. Psychol.*, 1960, **11**, 71–98.

Meyer, D. R. On the interaction of simultaneous responses. *Psychol. Bull.*, 1953, **50**, 204–220.

Miller, R. B. Handbook on training and training equipment design. *USAF WADC tech. Rep.*, 1953, No. 53–136.

Munn, N. L. Bilateral transfer of learning. *J. exp. Psychol.*, 1932, **15**, 343–353.

Naylor, J. C., & Briggs, G. E. Long-term retention of learned skills: A review of the literature. *USAF ASD Tech. Rep.*, 1961, No. 61-390.

Noble, C. E., & Alcock, W. T. Human delayed-reward learning with different lengths of task. *J. exp. Psychol.*, 1958, **56**, 407–412.

Noble, C. E., & Broussard, I. G. Effects of complex transformations of feedback upon simple instrumental behavior. *J. exp. Psychol.*, 1955, **50**, 381–386.

Noble, C. E., & Farese, F. J. An apparatus for research in human selective learning. *J. Psychol.*, 1955, **39**, 475–484.

Osgood, C. E. The similarity paradox in human learning: a resolution. *Psychol. Rev.*, 1949, **56**, 132–143.

Parker, J. F., Jr., & Fleishman, E. A. Use of analytical information concerning task requirements to increase the effectiveness of skill training. *J. appl. Psychol.*, 1961, **45**, 295–302.

Payne, B. Does the cumulative mean measure 'primary growth'? *J. exp. Psychol.*, 1942, **30**, 512–514.

Poulton, E. C. On prediction in skilled movements. *Psychol. Bull.*, 1957, **54**, 467–478.

Pyle, W. H. Economical learning. *J. educ. Psychol.*, 1913, **4**, 148–158.

Pyle, W. H. Concentrated versus distributed practice. *J. educ. Psychol.*, 1914, **5**, 247–251.

Pyle, W. H. Transfer and interference in card-distributing. *J. educ. Psychol.*, 1919, **10**, 107–110.

Renshaw, S., & Schwarzbek, W. C. The dependence of the form of the pursuitmeter learning function on the length of the inter-practice rests: I. Experimental. *J. gen. Psychol.*, 1938, **18**, 3–16.

Reynolds, B. The effect of learning on the predictability of psychomotor performance. *J. exp. Psychol.*, 1952, **44**, 189–198.

Reynolds, B., & Adams, J. A. Motor performance as a function of click reinforcement. *J. exp. Psychol.*, 1953, **45**, 315–320.

Reynolds, B., & Adams, J. A. Psychomotor performance as a function of initial level of ability. *Amer J. Psychol.*, 1954, **67**, 268–277.

Robinson, E. S. The 'similarity' factor in retroaction. *Amer. J. Psychol.*, 1927, **39**, 297–312.

Robinson, E. S. Work of the integrated organism. In C. Murchison (Ed.), *Handbook of general experimental psychology*. Worcester, Mass.: Clark Univer. Press, 1934.

Robinson, E. S., & Heron, W. T. The warming-up effect. *J. exp. Psychol.*, 1924, **7**, 81–97.

Rockway, M. R. Bilateral reminiscence in pursuit-rotor learning as a function of amount of first-hand practice and length of rest. *J. exp. Psychol.*, 1953, **46**, 337–344.

Rosenquist, H. S. The effect of active observation on pursuit-rotor performance. Unpublished doctoral dissertation, Tulane Univer., 1964.

Ruch, T. C. Factors influencing the relative economy of massed and distributed practice in learning. *Psychol. Rev.*, 1928, **35**, 19–45.

Ryan, F. J., & Bilodeau, E. A. Countertraining of a simple skill with immediate and 1-week delays of informative feedback. *J. exp. Psychol.*, 1962, **63**, 19–22.

Saufley, W. H., Jr., & Bilodeau, I. McD. Protective self-pacing during learning. *J. exp. Psychol.*, 1963, **66**, 596–600.

Sears, R. R., & Hovland, C. I. Experiments on motor conflict: II. Determination of mode of resolution by comparative strengths of conflicting responses. *J. exp. Psychol.*, 1941, **28**, 280–286.

Seashore, R. H., Buxton, C. E., & McCollom, I. N. Multiple factorial analysis of fine motor skills. *Amer. J. Psychol.*, 1940, **53**, 251–259.

Siipola, E. M. Implicit or partial reversion-errors: a technique of measurement and its relation to other measures of transfer. *J. exp. Psychol.*, 1940, **26**, 53–73.

Siipola, E. M., & Israel, H. E. Habit-interference as dependent upon stage of training. *Amer. J. Psychol.*, 1933, **45**, 205–227.

Silver, R. J. Effect of amount and distribution of warming-up activity on retention in motor learning. *J. exp. Psychol.*, 1952, **44**, 88–95.

Smith, K. U. The geometry of human motion and its neural foundations: I. Perceptual and motor adaptation to displaced vision. *Amer. J. phys. Med.*, 1961, **40**, 71–87.

Smith, K. U. *Delayed sensory feedback and behavior.* Philadelphia, Pa.: Saunders, 1962.

Smith, K. U., & Smith, W. M. *Perception and motion.* Philadelphia, Pa.: Saunders, 1962.

Smith, K. U., Wargo, L., Jones, R., & Smith, W. M. Delayed and space-displaced sensory feedback and learning. *Percept. mot. Skills*, 1963, **16**, 781–796.

Smith, K. U., with Gould, J., & Wargo, L. Spatial organization of neurobehavioral systems. *Amer. J. phys. Med.*, 1964, **43**, 49–84.

Snoddy, G. S. *Evidence for two opposed processes in mental growth.* Lancaster, Pa.: Science Press, 1935.

Starch, D. A demonstration of the trial and error method of learning. *Psychol. Bull.*, 1910, **7**, 20–23.

Starch, D. Periods of work in learning. *J. educ. Psychol.*, 1912, **3**, 209–213.

Stratton, G. M. Some preliminary experiments on vision without inversion of the retinal image. *Psychol. Rev.*, 1896, **3**, 611–617.

Stratton, G. M. Vision without inversion of the retinal image. *Psychol. Rev., 1897*, **4**, 341–360.

Swift, E. J. The acquisition of skill in type-writing; a contribution to the psychology of learning. *Psychol. Bull.*, 1904, **1**, 295–305.

Swift, E. J. Memory of a complex skillful act. *Amer. J. Psychol.*, 1905, **16**, 131–133.

Swift, E. J. Memory of skillful movements. *Psychol. Bull.*, 1906, **3**, 185–187.

Swift, E. J. Relearning a skillful act: An experimental study in neuro-muscular memory. *Psychol. Bull.*, 1910, **7**, 17–19.

Swift, E. J., & Schuyler, W. The learning process. *Psychol. Bull.*, 1907, 4, 307–310.

Taylor, A., & Noble, C. E. Acquisition and extinction phenomena in human trial-and-error learning under different schedules of reinforcing feedback. *Percept. mot. Skills*, 1962, **15**, 31–44.

Thorndike, E. L. *Animal intelligence*. New York: Macmillan, 1911.

Thorndike, E. L. *Educational psychology*. New York: Teachers College, Columbia Univer., 1914.

Thorndike, E. L. The law of effect. *Amer. J. Psychol.*, 1927, 39, 212–222.

Thorndike, E. L. *Human learning*. New York: Century, 1931.

Thorndike, E. L. *The fundamentals of learning*. New York: Teachers College, Columbia Univer., 1932.

Thorndike, E. L., & Woodworth, R. S. I: The influence of improvement in one mental function upon the efficiency of other functions. II: The estimation of magnitudes. III: Functions involving attention, observation and discrimination. *Psychol. Rev.*, 1901, 8, 247–261, 384–395, and 553–564.

Underwood, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49–60.

Underwood, B. J., & Postman, L. Extraexperimental sources of interference in forgetting. *Psychol. Rev.*, 1960, **67**, 73–95.

Walker, L. C., DeSoto, C. B., & Shelly, M. W. Rest and warm-up in bilateral transfer on a pursuit rotor task. *J. exp. Psychol.*, 1957, **53**, 394–398.

Waters, R. H., & Poole, G. B. The relative retention values of stylus and mental maze habits. *J. exp. Psychol.*, 1933, **16**, 429–434.

Wheeler, R. H., & Perkins, F. T. *Principles of mental development*. New York: Crowell, 1932.

Whipple, G. M. *Manual of mental and physical tests*. Baltimore, Md.: Warwick & York, 1910.

Wolfle, D. Training. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.

Woodrow, H. The relation between abilities and improvement with practice. *J. educ. Psychol.*, 1938, **29**, 215–230. (a)

Woodrow, H. The effect of practice on groups of different initial ability. *J. educ. Psychol.*, *1938*, **29**, 268–278. (b)

Woodrow, H. The effect of practice on test intercorrelations. *J. educ. Psychol.*, 1938, **29**, 561–572. (c)

Woodrow, H. Factors in improvement with practice. *J. Psychol.*, 1939, **7**, 55–70. (a)

Woodrow, H. The relation of verbal ability to improvement with practice in verbal tests. *J. educ. Psychol.*, 1939, **30**, 179–186. (b)

Woodrow, H. Interrelations of measures of learning. *J. Psychol.*, 1940, **10**, 49–73.

Woodrow, H. The ability to learn. *Psychol. Rev.*, 1946, **53**, 147–158.

Yarmolenko, A. The motor sphere of school-age children. *J. genet. Psychol.*, 1933, **42**, 298–318.

Zeaman, D., & Kaufman, H. Individual differences and theory in a motor learning task. *Psychol. Monogr.*, 1955, **69**, No. 6 (Whole No. 391).

# Chapter 2

# Selective Learning[1]

CLYDE E. NOBLE
*The University of Georgia*

Human behavior is predictable. This inductive generalization is as widely accepted by psychologists as it is doubted by laymen. Behavior is measurable in different ways and it is a function of many variables – some physical, others biological, still others social. Interactive all, these three sets of factors combine to produce in human nature one of the most complex and challenging problems scientists have ever attacked.

The topic of the present chapter is nonverbal selective learning: how man acquires the ability to produce an available motor response on cue as distinct from learning how to make a new response. Attention is focused on the simpler associative processes with emphasis on the Three Rs of Learning: Response, Repetition, Reinforcement.

## Multiple-Choice Situations

### THORNDIKE'S CONTRIBUTIONS

The 19th century origins of contemporary interest in multiple-choice learning are found among the behavior-oriented zoologists of the post-Darwinian era. Romanes, Morgan, Loeb, and Jennings were the leaders. Their mission: to undertake a comparative analysis of the intellectual and problem-solving abilities of infrahuman organisms. Hobbled by the limitations of naturalistic observation and the anecdotal method, however, early comparative psychology failed to accomplish much of scientific value until the advent of Thorndike (1898). He applied the experimental method to animal trial-and-error learning via the puzzle box. Thorndike's influential formulations of the empirical law of effect and the theory of connectionism (selecting and connecting) contributed significantly toward transforming classical association psychology into a genuine laboratory science. He observed that his hungry cats' behavior on the initial trials was characterized by errorful groping and extended latencies until they chanced to trip the release mechanism which provided access to the lure. With successive reinforcements the subjects' errors and latencies gradually diminished; eventually they came to restrict their movements to the critical area of the apparatus and thereby selected the correct or adaptive responses more efficiently.

---

[1] Portions of this chapter were delivered as invited addresses before colloquia at the University of Georgia, Florida State University, the University of Alabama, and Vanderbilt University during 1964–1965. For financial assistance I am indebted to Dean Robert A. McRorie, Director of the University's Office of General Research, and to Dr. Joseph C. Hammock, Head, Department of Psychology. My research team included Karen L. Buschow, C. Dwaine Cochran, Joan E. Eldodt, Judy H. Hardman, Helen R. Jones, Thomas A. Jones, and Ella M. Rowe. Helpful comments were received from Drs. M. H. Hodge, D. E. McAllister, W. R. McAllister, E. D. Neimark, and L. J. Peacock.

Thorndike saw no necessity for assuming any ideational processes to explain his cats' behavior—nor did Guthrie (1952) with more elaborate instrumentation—but those of a Gestalt turn of mind set about to design simpler problems which would permit animals to produce sudden ("insightful") solutions (Hilgard, 1956). Subsequent investigators explored a variety of string-pulling, bar-pressing, key-pecking, stick-pushing, and latch-opening tasks, with rewards running the gamut from food to shock-avoidance to poker chips. Although the trial-and-error vs. insight issue (a subdivision of the continuity-noncontinuity issue) has still not been resolved to everyone's satisfaction, the work of Harlow (1959) on learning-set formation in primates has convincingly demonstrated that either gradual or sudden acquisition curves can be produced through the manipulation of transfer-of-training variables.

### Yerkes' Multiple-Choice Method

Thorndikean apparatus was designed only to reflect gross $R_t$ (time) measures. No provision was made to record $R_f$ (frequency) scores or to isolate the varied behaviors of the subjects in puzzle boxes. It remained for Yerkes (1914) to contribute a more analytic technique by elaborating the animal quadruple-choice method of Hamilton (1911) into a generalized multiple-choice method for man. Originally developed for studying different species' stereotyped reactions to insoluble problems, the Hamilton technique was modified by Yerkes to provide five new emphases: (1) $R_f$ scores in addition to $R_t$ scores, (2) acquisition of correct responses rather than extinction of erroneous responses, (3) variations in complexity, (4) soluble problems instead of insoluble problems, (5) wider interspecies applicability. For human subjects Yerkes' (1921) apparatus contained 12 retractable keys, two or more of which would be presented on a given trial (key setting). The subject's task was to avoid the incorrect keys (R−) and press the correct key (R+). Reinforcing feedback was provided by a buzzer following each R+, and practice was continued until either mastery or some arbitrary proficiency criterion was reached. Yerkes favored relational problems (e.g., middle key of five) whose elements were changed from trial to trial. He believed that their solutions entailed ideational or reasoning abilities based upon the "perception of a certain constant relation among a series of objects" (Yerkes, 1916, p. 10).

A psychologist who called himself a Behaviorist (Yerkes, 1914) yet espoused many Gestalt principles, Yerkes is perhaps best considered a Functionalist. He was clearly less interested in the multiple-choice method as a reference task for discovering the laws of selective learning than in its utility for testing animal intelligence. Thus, he avoided studies of serial learning, retention, and transfer of training. Had he not been so reluctant to employ the Yerkes Box as a learning device he might have found, as I did (Noble, 1957a), that it is a fairly reliable piece of apparatus when used in a deliberately unconventional manner. This was achieved in our laboratory by defining settings as choice points instead of trials. We gave his most difficult problem (Yerkes, 1921, No. 10) to 100 college subjects using a commercial Yerkes Box (by Stoelting) where the solution was to alternate left and right of any middle key. Presenting just four trials of 10 choices each, an odd-even-trials reliability coefficient (corrected) of .85 was obtained for $R_f$ scores.

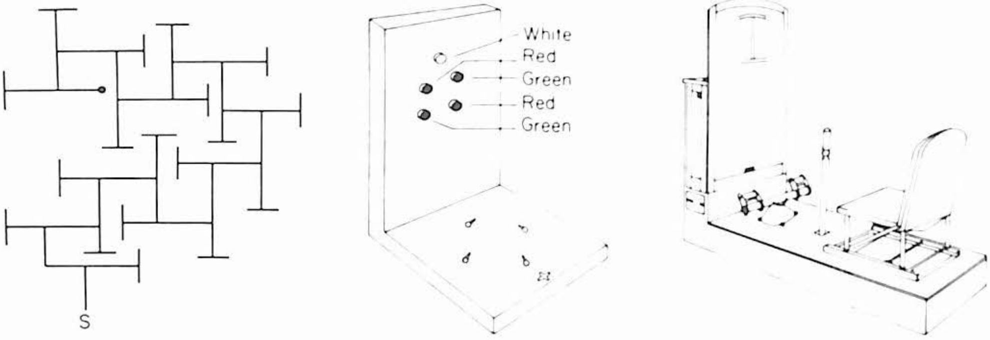
Unfortunately, as compared with the fruitfulness of the classical-conditioning, maze-learning, operant-response, and discrimination-learning techniques, little data of any systematic importance had emerged from this method by the 1930s when learning theories came into vogue.[2] The Yerkes Box was generally bypassed in favor of Köhler's tool-using situations, Hunter's spatial and temporal mazes, Maier's reasoning tasks, the Lashley jumping stand, Tolman's T mazes, the Skinner Box, Spence's discrimination problems, the Mowrer-Miller Box, and Harlow's Wisconsin General Test Apparatus.

## CONTEMPORARY HUMAN APPARATUS

In Fig. 1 are line drawings of the selective-learning devices most frequently cited in research reports of recent vintage. The paragraphs below

---

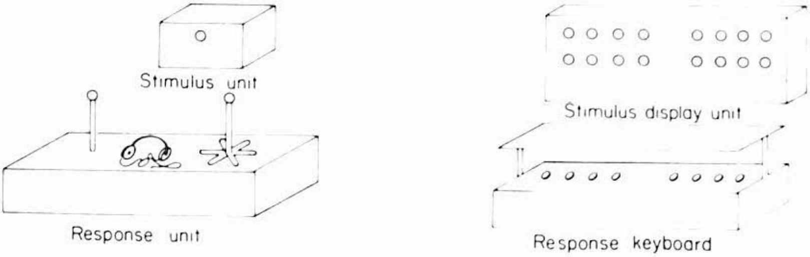
[2] To the end of his life Yerkes remained optimistic about the untapped resources of the multiple-choice method; he encouraged younger psychologists to exploit its potentialities (personal communication, 1954). My own emphasis upon measuring relative frequency of $R+$, i.e. empirical response probability ($R_p$) scores, on the Selective Mathometer (cf. Fig. 1), may be regarded as a further development of the Thorndike-Yerkes tradition, coupled with contemporary experimental and statistical methodology. How this technique can be integrated into theoretical psychology is taken up in the section on Theory of Selective Learning.



Finger maze

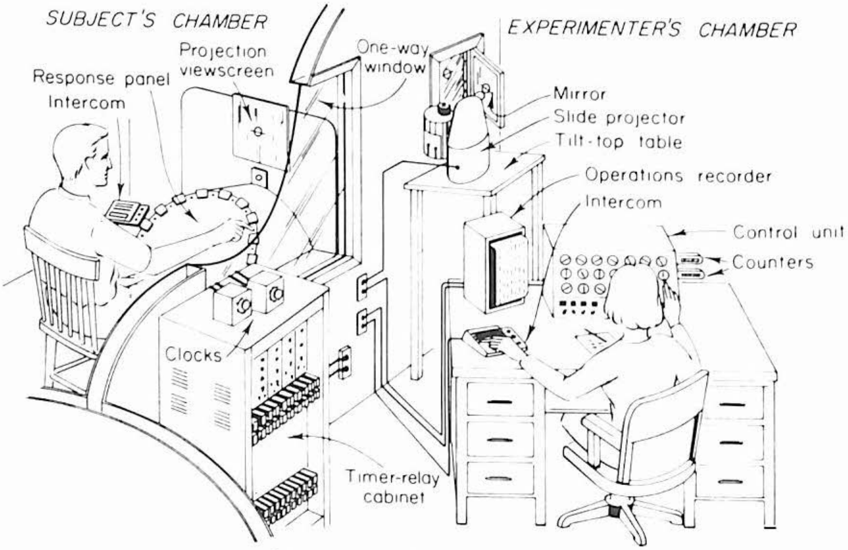
Discrimination reaction timer

Complex coordinator



Star discrimeter

Multiple serial discrimeter



Selective mathometer

FIG. 1. Apparatus for research in human selective learning.

contain brief summaries of the functions and reliabilities of each apparatus. More detailed descriptions, wiring diagrams, etc. may be found in the original sources.

Oldest of all apparatus for trial-and-error learning is probably the human maze. The great hedge-maze of Henry VIII's Hampton Court Palace in England served as the prototype for the first rat maze of 1899. After several simplifications, the high-relief Finger Maze, designed by Walter R. Miles in 1928, became the most popular human laboratory version (Woodworth & Schlosberg, 1954). The blindfolded subject's task is to trace the pattern with his forefinger from start to goal as quickly as possible and with a minimum of errors. Proficiency is typically measured by $R_t$ and $R_r$ scores, employing the correction procedure. Three different forms (T, U, Y) and many lengths (10 to 40 choice points) have been used, cognate varieties being "mental" (verbal), stylus, punchboard, and bolthead mazes. Since the latter two possess the virtue of allowing either the correction or the noncorrection procedure, extra feedback stimuli (usually auditory signals) have to be added to the tactile and kinesthetic cues normally present for reinforcement. Reliability coefficients of .927 and .955 have been reported for stylus and punchboard error scores, respectively (Hilgard, 1951); coefficients varying from .600 to .975 for different methods and groups on Miles' Finger Maze (Spence, 1932); and coefficients of .772 for errors and .967 for speed scores on a semilinear multiple-T Finger Maze of the Warden type (Noble, Noble, & Alcock, 1958c).

Next in seniority is the USAF Complex Coordinator, a World War II modification of a cockpit-simulating apparatus designed by Neely C. Mashburn in 1934 (Melton, 1947). The subject's task is to learn to align a pattern of red lights presented in three dimensions with a matching pattern of green lights. This is done by appropriate movements of the handstick (aileron, elevator) and footbar (rudder) controls. When the lights have been correctly matched for .5 second a new pattern appears automatically, conferring reinforcing feedback. A modified-correction method is the only one available. Proficiency is typically measured by cumulating the subject's $R_r$ scores over a block of 8-minute trials, although time per match has been used also. Reliability coefficients range from .710 to .917 depending upon the correlation method, sample size, and stimulus patterns chosen (Melton, 1947; Noble, 1952a,b,c).

Another World War II product of the USAF is the Discrimination Reaction Timer, designed by John T. Finan in 1942 (Melton, 1947). The subject's task is to learn to snap one of four toggle switches in response to the simultaneous onset of a pair of red and green signal lamps. The relative position of the red lamp with respect to the green determines which of the four switches is correct. A modified-correction procedure is the only one available. Reinforcing feedback is provided by the offset of a white light. Proficiency is typically measured by cumulating a subject's $R_t$ scores over a trial block of 20 3-second stimulus presentations, with an interblock interval of 40 seconds. Reliability coefficients vary from .870 to .930 (Melton, 1947).

The Star Discrimeter was designed by Carl P. Duncan and Benton J. Underwood in 1950 under a contract with the USAF. The subject's task is to learn to respond to each of six colored lights emanating from the stimulus unit by moving the right-hand lever into one of six corresponding slots. The slots are arranged radially in the shape of an asterisk on the response unit, while the colors appear automatically in a continuous random sequence of red, orange, and yellow hues. The task is usually performed while executing a hand-steadiness action with the left hand, error information being delivered by earphones. The correction method is the only one available, reinforcing feedback occurring when a new light comes on. Proficiency is typically measured by computing a subject's $R_t$ scores during a 20-second trial. Fixed intertrial intervals of 10 seconds are conventional, but interblock intervals have varied from 1 to 3 minutes. Reliability coefficients range from .883 to .965 (Duncan, 1953). The Iowa model (McAllister, 1953) omits the steadiness test.

The Multiple Serial Discrimeter was designed by David A. Grant in 1951 under a contract with the USAF. The subject's task is to learn to respond to two of eight red stimulus lights on a horizontal display by pressing the correct two of eight reaction keys on the control panel below. This matches up a pair of green lamps to provide reinforcing feedback. Two hands are employed by the subject, the left operating keys 1–4, the right operating keys 5–8. Any number of lights can be presented, and the light-key relationships may be scrambled. Either apparatus pacing or self pacing is possible, but the latter (correction) method is most common. A trial consists of a presentation of two stimulus lights, which are grouped randomly in blocks of 25 with a 1-minute rest between blocks. Proficiency is typically measured by computing $R_t$ scores per trial, but R− scores are also used. Reliability coefficients have not been reported (Morin & Grant, 1955).

The Selective Mathometer, designed by Clyde E. Noble in 1952, is the fifth legacy of the USAF psychology program to the field of selective learning (Noble & Farese, 1955), and it is now fully automatic (Noble, Fuchs, & Thompson, 1963). Figure 1 shows the present layout in the Georgia laboratory. The subject's task is to discover which one of 19 microswitch keys (numbered left to right) is to be pressed in response to each of a series of distinctive slides (e.g., circuit symbols) presented on the screen. The keys are arranged in a semicircle on the reaction panel. Correction, modified-correction, and noncorrection methods are available, and there is separate control of stimulus duration, interstimulus interval, intertrial interval, and stimulus sequence. The reinforcing stimulus ($S_g$) is the onset of a green light centered above the keyboard. Either serial or terminal $S_g$ is delivered automatically, together with independent variation of delay, duration, and periodicity. All events are recorded by counters and by a continuous graphic recorder. Proficiency is typically measured by computing a subject's $R_p$ scores over a trial of 10 different 2-second stimulus presentations, although $R_t$ scores have been used also. The interstimulus and intertrial intervals are usually 2 seconds and 8 seconds, respectively. Reliability coefficients range from .938 to .989, varying with type of score, training method, and task complexity (Noble, 1957a, 1965a; Noble *et al.*, 1958c).

The foregoing does not exhaust the list of selective-learning apparatus. Other devices with commendable features are those designed by Adams (1954a,b,c), Barch (1960b), Besnard, Briggs, Mursch, & Walker (1955), Bilodeau (1953), Birch (1964), Mandler (1954), Seibel (1962), Seward, Dill, & Holland (1944), and Spiker and Holton (1958). In addition there are discrimination-learning tasks and double-bar Skinner Boxes of varied design.

### VERBAL-EXPECTATION TASKS

One might wonder why I have not included some of the apparatus used in symmetrical-choice experiments of the Humphreys (1939) type. These have variously been called "verbal expectation," "event prediction," "verbal conditioning," "light guessing," "probability discrimination," and "probability learning" studies. Subjects are required to "guess" (predict, anticipate) which of two alternative events, randomly varying in predetermined relative frequencies, will occur on a given trial. It happens that their choices are often more in accord with the actual proportions of occurrence than with a simple accuracy-maximizing rule (e.g., always predict the more frequently occurring event). This phenomenon (called "probability matching"), together with other characteristics, has encouraged the development of several ingenious statistical or stochastic learning models (Bush & Estes, 1959; Bush & Mosteller, 1955; Estes, 1959, 1962, 1964; Restle, 1961).

Despite the importance of verbal-expectation research for the testing of mathematical models, I am not convinced that it represents an instance of selective learning. It is at least an exotic kind of learning situation to manifest the following half-dozen anomalies: (1) $R_p$ scores in acquisition and extinction which are unaffected by intertrial intervals varying from 5 seconds to 40 seconds (Grant, Hornseth, & Hake, 1950); (2) $R_t$ scores which are independent of practice and uncorrelated with $R_p$ scores (Gerjuoy, Gerjuoy, & Mathias, 1964); (3) acquisition rates which do not depend on initial conditions or schedules of reinforcement (Estes & Straughan, 1954); (4) equivocal effects on performance when amount of reward is varied (Bush & Mosteller, 1955; Estes, 1962); (5) failure of probability matching when "guessing" is made secondary to motor skill (Wyckoff & Sidowski, 1955); (6) occurrence of probability matching for several hundred trials, but no evidence beyond that up to 1000 trials (Edwards, 1961).

Furthermore, there are certain aspects of the theoretical models which pose difficulties in applying them to traditional classical or instrumental conditioning situations. These include an adherence to the Markovian assumptions of path independence and constant transitional probabilities (Bower, 1959; Bush & Mosteller, 1955; Estes, 1959); an insensitivity to memory and transfer effects in two-choice noncontingent schedules (Anderson, 1960, 1964); an apparent insouciance with respect to individual differences (Restle, 1961); and a failure to predict inflection points below $R_p = .50$ in conditioning and selective-learning experiments. Even data from the Humphreys task itself are not always predictable from the regnant models (Estes, 1964; Witte, 1964). This combination of experimental and theoretical eccentricities has produced in some psychologists feelings of diffidence about the implications of verbal-expectation research (Anderson & Grant, 1957; Anderson, 1964; Hake, Grant, & Hornseth, 1951; Kendler, 1959; Spence, 1960; Underwood, 1953).

My fundamental objection is that selective-learning tasks are typically situations where only one R+ is reinforced 100% of the time in the presence of a particular S. In mathometry we never divide the reinforcement within a trial between Rs for a given stimulus, as in the Humphreys task. It is true that in habit-reversal investigations we alternate reinforcement, but the partiality of the reinforcement is between blocks, not within blocks. The verbal-expectation experiment is one in which successive acquisition and extinction processes are produced by means of a random pattern of partial reinforcement. Of course, the selective-learning situation becomes a verbal-expectation task if the stimuli are not discriminable from one choice point to the next. Perhaps it is instructive to regard verbal-expectation learning as the limiting case at the transition between discrimination learning and concept formation.

A growing conviction among the younger math-modeleers (being privately voiced more frequently today) is that statistical learning theories seem to be too simplified for the tasks the subjects are required to perform, and that some of the reference experiments have generated empirical phenomena which may be atypical. Because I believe that the laws and theories of selective learning ought to be continuous with those of classical and instrumental conditioning (besides showing a high degree of internal consistency), I shall not attempt to subsume data from verbal-expectation tasks under the present rubric.

## Descriptive Analysis of Selective Learning

By a selective-learning task I refer to any multiple-choice situation in which an organism acquires the ability (1) to make a joint selection from the relevant stimuli present and from its own repertoire of available discrete responses, (2) to form the necessary associative connections between the stimuli and responses required by the task, and (3) to link these S-R pairs together in a coordinated behavior sequence. The emphasis is upon cuing, pairing, and chaining rather than upon the development of fine perceptual discriminations or precision motor skills. Selective-learning tasks, in short, deemphasize perceptual learning as well as response learning. Rarely does selective learning require gradations in amplitude, force, or timing of action as do the classical perceptual-motor skill devices employed for research in tracking, dexterity, and compensatory pursuit (Adams, 1964; Bilodeau & Bilodeau, 1961; Fitts, 1964).

Of course, not all selective-learning tasks are factorially pure in this respect, but the devices picture in Fig. 1 are representative of my present emphasis. Fitting the criteria best, I believe, are the Multiple Serial Discrimeter (MSD) and the Selective Mathometer (SM). The Discrimination Reaction Timer (DRT) is weighted somewhat heavily on the perceptual side (Fleishman & Hempel, 1955), just as the Complex Coordinator (CC) leans toward the motor side (Fleishman & Hempel, 1954; Noble, 1952c, 1953b). The Finger Maze (FM) and the Star Discrimeter (SD) fall somewhere in between, to the left and right, respectively. Thus, the perceptual-motor continuum of my mind's eye has the approximate order sketched below:

| DRT | FM | MSD | SM | SD | CC |
|---|---|---|---|---|---|
| Perceptual emphasis | | Selective learning emphasis | | | Motor emphasis |

There are four basic elements in selective-learning situations (Spence, 1956): a motivated organism, a multiple-choice stimulus situation in the external environment, the organism's initial repertoire of discrete responses, and the subsequent physical events together with their after-effects which are relevant to the organism's motivational state. In Fig. 2 is a diagram of a simple selective-learning problem in which all of the dominant responses($R_1$, $R_2$, $R_3$) are irrelevant or *incorrect* (R−). The percentage values above the arrows denote reaction tendencies of different relative initial strengths. Arbitrarily I have designated the last response ($R_n$) as the *correct* one (R+); i.e., the first-to-be-learned pair is $S_n$ − − −> $R_n$. This paradigm illustrates the fact that organisms do not enter upon multiple-choice situations behaving in a "blind" or "random" fashion, but rather they arrive with a hierarchy of systematic response tendencies which, on the basis of hereditary or acquired factors, are more likely to be uncorrelated with the task designed by the experimenter than to be correlated positively or negatively (Hull, 1930, 1952). Thus it

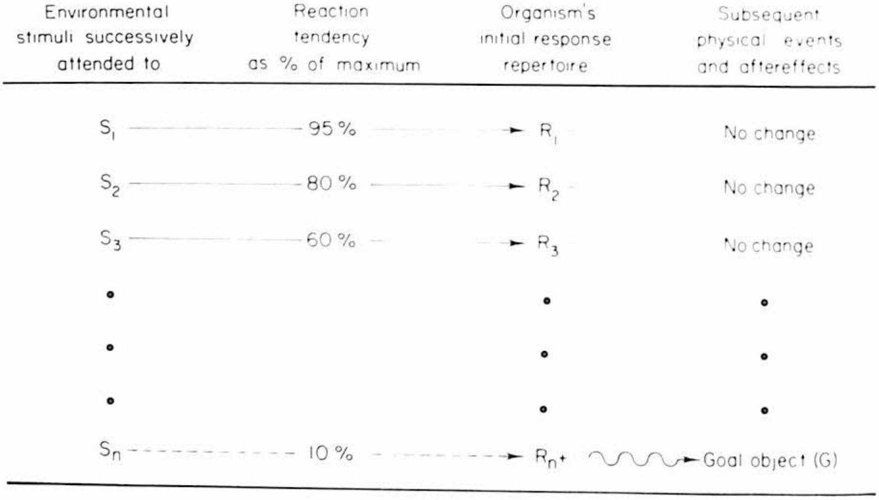

Environmental stimuli successively attended to
Reaction tendency as % of maximum
Organism's initial response repertoire
Subsequent physical events and aftereffects
S1
95 %
R1
No change
S2
80 %
R2
No change
S3
60 %
R3
No change
Sn
10 %
Rn+
Goal object (G)

FIG. 2. A simple (nonserial) selective learning situation.

comes about that the subject's initial habit hierarchy must be reorganized if he is to solve the problem. Selective learning *is* this reorganization.

By the law of effect and the S-R reinforcement theory, the probability, speed, and amplitude of R+ will be strengthened because the goal object (G) which follows it reinforces the habit connection. In the case of human subjects pressing buttons for green lights (cf. Fig. 1) the reinforcers are secondary; or, if you prefer, the feedback signals are informative. At any rate, learning is complete when errors (R−) are eliminated and correct responses (R+) occur promptly on every trial. For convenience and *faute de mieux*, I shall continue to refer to the class of goal events (mostly green lights) by the symbol $S_g$.

After description, explanation. In the next section we undertake a theoretical analysis of multiple-choice phenomena.

## Theory of Selective Learning

### ROLE OF THEORY

It is sometimes necessary to remind ourselves that a psychological theory is not intended to describe the *structure* of any particular living being. Rather, it is designed to provide a model—a schematic, hypothetical organism—whose *functions* are analogous to those of real organisms. A theory's chief aim is thus to explain and predict a large number of phenomena with a small number of assumptions. Obviously, then, theoretical models are not supposed to be mere collections of empirical facts. They are broad conceptual paradigms which permit scientists to arrange facts into meaningful patterns. The principal criteria of a valid theory are two in number: (1) it must provide the basis of interpreting or understanding inductively-known facts, and (2) it should aid deductively in the discovery of new facts. Whether a theory is statistical, topological, mechanical, chemical, or neural is beside the point. What matters is that it have the capability of explanation and prediction.

Roughly intermediate in complexity between classical conditioning and perceptual-motor skill acquisition, selective learning presents formidable theoretical obstacles because of the multiplicity of events occurring, the inadequacy of experimental control over the relevant variables, and, as we shall see, some mathematical quirks. When R+, R−, and $R_t$ scores are recorded concurrently in analytic multiple-choice tasks like the Selective Mathometer (cf. Fig. 1), it is clear that several different processes are going on at about the same time: (1) strengthening of R+ through reinforcement, (2) weakening of R− through nonreinforcement, (3) occasional recovery of an extinguished R− with time, (4) generalization of approach from R+ to R−, (5) generalization of avoidance from R− to R+. The isolation and explanation of these processes require clever experimental design combined with ingenious quantitative theorizing via intervening variables.

## THE HULL-SPENCE THEORY

Hull's (1930) grand strategy was to construct a theory of complex learning by building upon the foundation of simple principles discovered in classical and instrumental conditioning experiments. Once the foundation was laid (Hull, 1943) he planned to return to his earlier interest in selective learning. In his last work Hull (1952) distinguished between problems of *response* selection (trial-and-error learning) and problems of *stimulus* selection (discrimination learning). Using the multiple-T Finger Maze as an illustration, trial-and-error learning would be required if the subject must respond invariably to the same side of the T (disregarding the width of the maze's two arms), whereas discrimination learning would be required if consistent response must be made to the wider arm (avoiding the narrow arm) regardless of position. Hull further differentiated between *simple* (single choice point) and *compound* (serial choice point) selective-learning situations. Finally, when multiple stimuli are combined with competing responses in the same task, he suggested that we speak of *mixed* selective learning. In view of the fact that most of the tasks considered in this chapter involve the development of adaptive behavior in making all-or-none choices among discrete alternatives on both the S and R sides, it is rarely necessary to use the qualifier. For my purpose the term *selective learning* nicely emphasizes the Darwinian as well as the probabilistic connotations.

Spence (1960, pp. 366–392) has recently reviewed several theoretical paradigms of selective learning favored by Hullians, including competing-S and competing-R models of spatial selective learning, S-type and R-type models of nonspatial selective learning, and compound S-R models for each.[3] The continuity principle is assumed to hold for all six paradigms. As an illustration of Spence's approach let us consider his S-R model for spatial selective (trial-and-error) learning. For reference, imagine a two-key task on the Selective Mathometer in Fig. 1 (cf. Spence, 1960, p. 380).

---

[3] Spence's classification into *spatial* and *nonspatial* selective-learning apparatus for animals is similar to Hull's given above; e.g., the dual-lever Skinner Box is a re-

At the choice point ($S_c$) the subject may respond either with a left-orienting ($r_L$) or with a right-orienting ($r_R$) receptor act, thus fixating and receiving exteroceptive stimulation from the visual discriminanda on the left ($S_{VL}$) or on the right ($S_{VR}$) in addition to differential kinesthetic stimuli ($s_{KL}$ or $s_{KR}$). If we assume that these orienting responses are mutually exclusive (incompatible), the probability of the occurrence of one or the other depends theoretically upon the difference between their reaction tendencies ($E$), modified by the Spencean postulates regarding the reaction limen ($L$) and oscillatory inhibition ($I_o$).[4] The concept of momentary reaction tendency ($\dot{E}$) is defined by the relation $\dot{E} = E - I_o$, where $I_o$ is a normally distributed decremental quantity varying randomly in magnitude over time. Momentary suprathreshold reaction tendency ($\dot{E}_{E-L}$) is specified in turn by the relation $\dot{E}_{E-L} = \dot{E} - L$, where $L$ is a hypothetical minimum value of $E$ necessary to produce a response. Now, as the number of trials ($N$) on which the correct orienting response ($r_R$) gets reinforced ($S_r$) is increased indefinitely, the difference between the $\dot{E}_{E-L}$ values of $r_R$ and $r_L$ becomes progressively greater, thereby raising the probability $P_1$ of $r_R$ (and lowering the probability $P_2$ of $r_L$) at that choice point. The latter inverse relationship follows from the fact that $P_1 + P_2 = 1.00$.

Similarly, the probabilities $P_3$ and $P_4$ depend on the extent to which the $\dot{E}$ values of their respective joint conditioned stimuli ($S_{VR} + s_{KR}$ vs. $S_{VL} + s_{KL}$) are suprathreshold as well as upon the difference between the two $\dot{E}$ values. Assuming that the two $I_o$ distributions are uncorrelated, have equal standard deviations ($\sigma_o$), and are constant over trials, conventional statistical tests of significance may be employed to compute the probabilities from the critical ratio formula $(\dot{E}_R - \dot{E}_L)/\sigma_d$ that the response of greater $\dot{E}$ will occur on successive trials. If all four $\dot{E}$ values were known, then on a given trial Spence would predict the probability that the subject will complete the response sequence leading to an R+ by the equation

$$R_p = \frac{P_3(1 - P_2P_4)}{P_3 + P_4 - P_3P_4} \tag{1}$$

A mathematical impasse arises at this point because in principle it is impossible to proceed deductively through the definitions of $\dot{E}$ and $\dot{E}_{E-L}$ directly to $R_p$ scores in Eq. (1) if sizable proportions of the two pairs of $\dot{E}$ distributions are subthreshold during the early trials. The necessity of evaluating double normal probability integrals can, however, be avoided by an iterative (approximation) technique which uses tabled values of the cumulative normal curve and deals only with the values of the final approach responses (Spence, 1956). As a matter of fact, even in the case of a single

---

sponse (spatial) selection device while the Lashley jumping stand is a stimulus (nonspatial) selection device.

---

[4] Brief accounts of the derivation of $R_p$ scores as a function of $N$ were given by Hull (1943, 1952) and Spence (1956, 1960). Lewis (1960) has provided a detailed computational example. For simplicity, I have omitted consideration of $I_n$ and $K$; hence we may dispense with the concept of effective reaction tendency $\bar{E}$ and assume that incentive is constant.

dominant reaction tendency, such as would arise in a classical conditioning experiment, a deductive barrier is encountered. Let me demonstrate.

Given $E = D \times H$ and substituting $H = M_H(1 - e^{-kN})$, we have

$$E = D[M_H(1 - e^{-kN})] \tag{2}$$

Since $\dot{E}_{E>L} = \dot{E} - L$, it follows from the oscillatory inhibition principle that $R_p$ is proportional to an integrated Gaussian probability function of the quantity $E_{E-L} > 0$; thus by the calculus

$$R_p = \int_{-2.5z}^{\dot{E}-L} \frac{1}{\sqrt{2\pi}}\, e^{\left(\frac{-z^2}{2}\right)}\, dz \tag{3}$$

where $z = x/\sigma$ transformation of the $I_o$ scale, and where the lower limit of integration is taken for convenience in order to deal with truncated distributions extending from the mean of $I_o \pm 2.5z$. Unfortunately, the distribution function postulated for $I_o$ by Spence's theory cannot be integrated to provide a simple equation in closed form for various values of $\dot{E}$ as $N$ increases without bound. Consequently, one would either have to derive an open-ended equation of the polynomial type (e.g., the Taylor series expansion)[5] or be content with a numerical approximation. Both Hull (1952) and Spence (1956) chose the second alternative, presumably because of the nonrational nature of the first alternative and the comparative ease of handling the normal curve tables.

### AN APPROXIMATION FORMULA

A third alternative is available. This consists of choosing a descriptive equation whose form and parameters can be rationalized and thereby coordinated with the theory. To be satisfactory the candidate selected should conform to most of the following desiderata: it should (1) bear close similarity to the numerical approximation of Eq. (3), (2) provide an excellent statistical fit to empirical $R_p$ data from conditioning and selective-learning experiments, (3) exist in closed form, (4) have fewer total parameters than Eqs. (2) and (3), (5) be capable of at least some parameter-free (pre-experimental) predictions consistent with the theory, and (6) be amenable to rapid curve-fitting techniques.

An equation satisfying all six criteria is the complex growth curve originated by Gompertz (1825) and elaborated by Courtis (1932), which I have adopted (Noble, 1954, 1957a,b) with the following psychological rationale:

---

[5] For example, the power series $R_p = k_o + k_1N + k_2N^2 + k_3N^3 + \ldots + k_nN^n$ contains linear, quadratic, cubic, etc. components up to the $n$th power if necessary.

$$R_p = a(i)^{r^N}, \tag{4}$$

where $R_p$ = probability of correct choice based on the pooled responses of a group of subjects, varying from .00 to 1.00; $a$ = asymptote or limit of $R_p$, theoretically dependent on feedback $(S_g)$ factors and possibly age $(O_y)$, sex $(O_s)$, and capacity $(O_c)$; $i$ = initial probability of R+ at outset of learning, given empirically by the reciprocal of the number of available independent and mutually exclusive responses and theoretically related to complexity $(S_c)$ and transfer $(S_t)$ factors; $r$ = rate parameter calculated by curve-fitting methods, theoretically governed by feedback factors $(S_g)$, complexity $(S_c)$, transfer $(S_t)$, work distribution $(S_w)$, aptitude $(O_h)$, and possibly age $(O_y)$, sex $(O_s)$, and motivation $(O_d)$; $N$ = number of practice trials providing differential reinforcement by correction, noncorrection, or modified-correction procedures.

Equation (4) specifies the dependent variable $(R_p)$ as a double exponential function of the major independent variable $(N)$; its three parameters govern the asymptote $(a)$, origin $(i)$, and rate of change $(r)$. With respect to $a$, if a given problem is soluble then $R_p$ will approach 1.00 as a function of successive reinforced trials. With respect to $i$, when no practice has been given, $N = 0$. Therefore, since $r^N = r^0 = 1.00$, the $R_p$ value at Trial $0 = a(i) = i$, which is the probability of an R+ occurring before the first choice point of Trial 1. For simplicity, I have disregarded the effects of transfer and position preferences. The former are expected to affect $r$, while the influence of the latter on $i$ is minimized by statistical counterbalancing of the possible response permutations (Noble, 1955, 1957a). By systematic variation of four stimulus variables $(S_c, S_g, S_t, S_w)$ and four organismic parameters $(O_d, O_h, O_s, O_y)$, holding constant the values of $a$ and $i$, it should be possible to rationalize $r$ by thus anchoring it to selected $S$ and $O$ factors.

In general, the curve of Eq. (4) exhibits a positive slope, is asymmetrical and sigmoidal, and has an inflection point (for soluble problems) at the ordinate $1/e = 1/(2.718) = .368$. So if $i < .368$, and if $i$ and $r$ vary between 0 and 1, it can be shown that the curve inflects at an abscissa value of $N > 0$. By setting the second derivative of Eq. (4) equal to zero and solving for $N$, we have

$$N = \frac{-\log_e(-\log_e i)}{\log_e r} \tag{5}$$

hence

$$\log_e r = \frac{-\log_e(-\log_e i)}{N} \tag{6}$$

Courtis (1932, p. 89) pointed out that a log-log transformation of the dependent variable is rectilinear over time ($N$ in our case), implying that the percentage of "growth" ($R_p$ in our case) increases to equal powers of itself in equal time periods ("isochrons"). He also worked out a molar theory which is remarkably similar to Hull's, even to the extent of incorporating postulates for variance and threshold. Later, a molecular derivation of Eq. (4) in terms of cellular growth was proposed by Shock and Morales (1942).

Equation (6) provides a convenient estimation formula for obtaining $r$ (Noble, 1957c). Consider a typical experiment in which subjects are practicing on the Selective Mathometer with three or more keys available. Normally they are either trained to a mastery criterion (five successive perfect trials), or a constant-trials design is used in which there are rational grounds for assuming $a = 1.00$. A free-hand curve is first drawn among the resulting data points from $i$ up to $R_p = .50$. Next, from the intersection of the fitted segment with an extended ordinate line at $R_p = .368$, a perpendicular is dropped to the abscissa. This $N_{\text{est}}$ value is substituted in Eq. (6), which is then solved for $r_{\text{est}}$ by taking the antilogarithm of the right-hand expression. Successive solutions of Eq. (4) will generate a set of theoretical (expected) response probabilities ($\tilde{R}_p$) which may then be compared with the empirical (observed) $R_p$ values for goodness of fit. By repeating the above procedure for selected values of $r_{\text{est}}$ (rounded to three decimal places), each time calculating the variance of the residual or vertical deviations ($\sigma^2_{\text{res}} = \Sigma[R_p - \tilde{R}_p]^2/n$), a least-squares solution may be obtained by minimizing $\sigma^2_{\text{res}}$. This method avoids the linear reduction procedure (Courtis, 1932; Lewis, 1960) which requires double logarithms, is more laborious, and gives a best fit only to the transformed data.

Finally, a goodness-of-fit index ($F$), expressed in percentage terms, is computed directly from the $\sigma^2_{\text{res}}$ value and the observed variance of the $R_p$ scores ($\sigma^2_{\text{obs}}$) by the formula

$$F = 100\left[1 - \frac{\sigma^2_{\text{res}}}{\sigma^2_{\text{obs}}}\right] \tag{7}$$

With Eq. (7), which is analogous to the coefficient of determination in linear regression, we may specify the percentage of variance in the dependent variable ($R_p$) accounted for by Eq. (4); i.e., $F$ measures the precision with which $R_p$ can be predicted from $N$ when $a$, $i$, and $r$ are constant.

In the case of experiments involving less than three choices or, more generally, when $i \geqslant 1/e$, the curve is negatively accelerated throughout. The fact that data from two-choice situations (e.g., most discrimination-learning tasks) rarely manifest an inflection point can, therefore, be nicely repre-

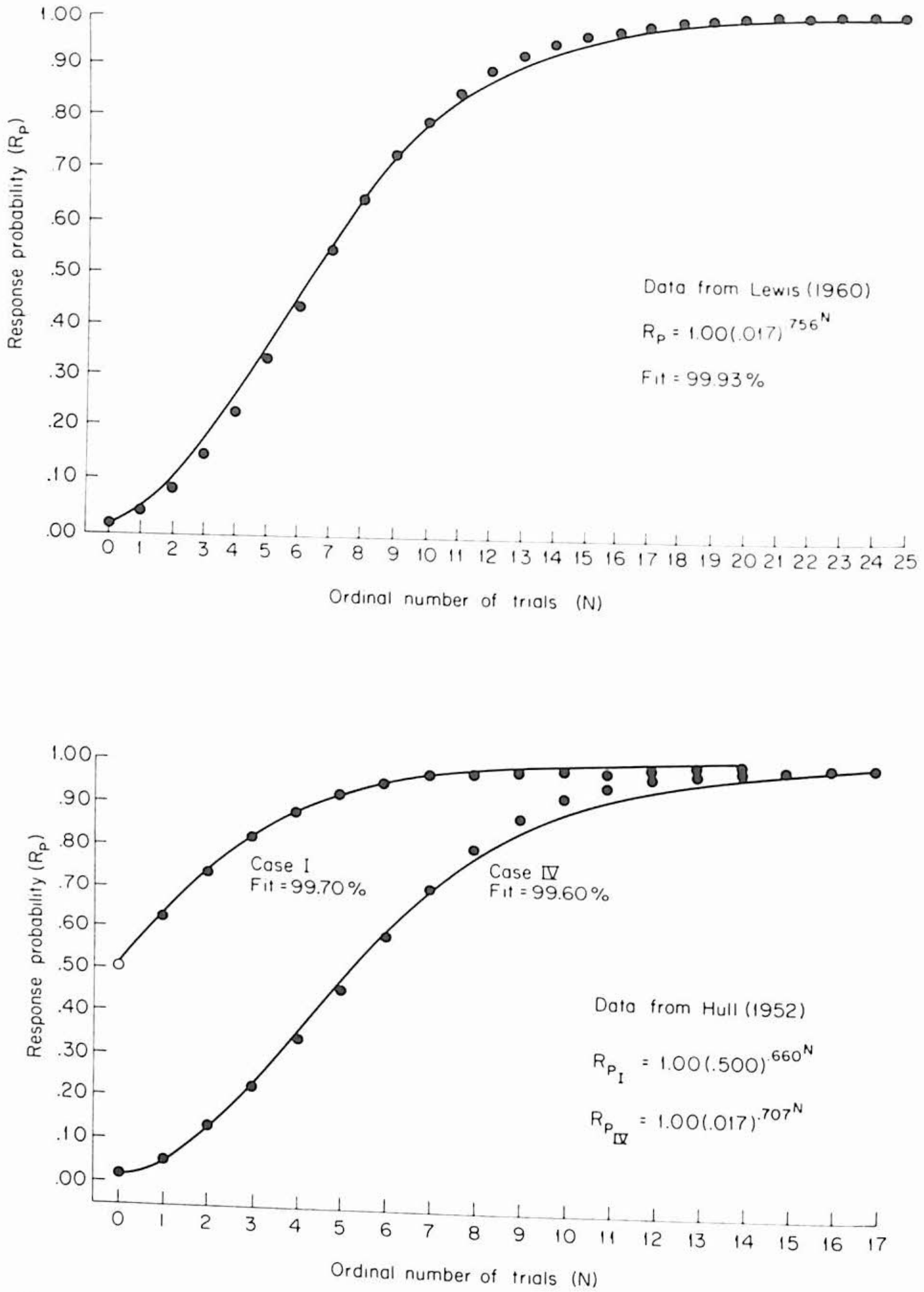
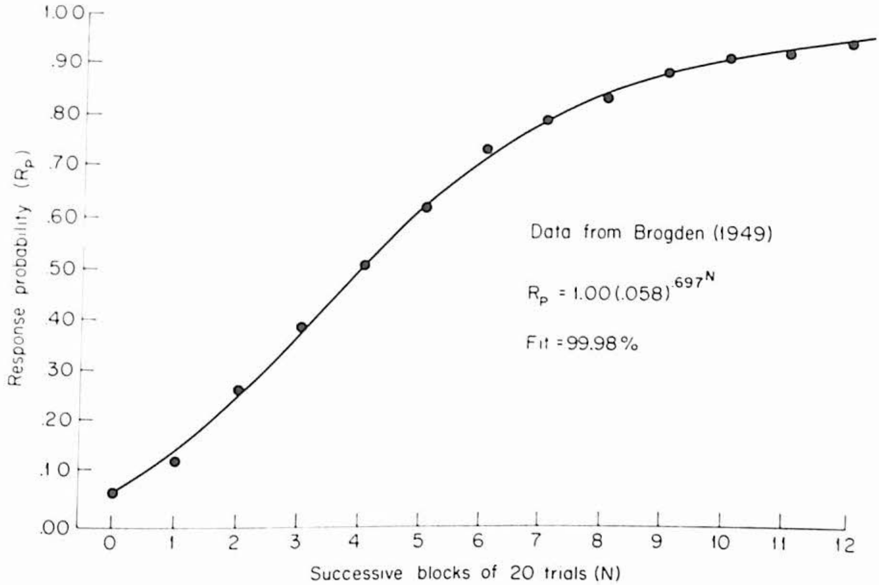
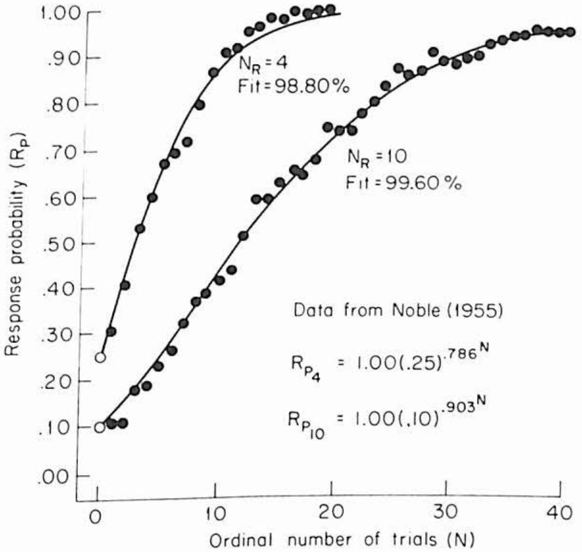


FIG. 3. Theoretical and empirical curves of response probability ($R_p$) as a function of number of trials ($N$), all fitted by Eq. (4). The two upper graphs are for conditioned-





response learning (after Lewis, 1960; Brogden, 1949). The two lower graphs are for selective learning (after Hull, 1952; Noble, 1955). Open circles denote rational origins.

Disagreeing with Peterson's interpretation, especially with its implications for Guthrie's (1952) recency principle, Voeks (1948) reanalyzed his data to find (1) that 68.4% of his 4518 recency predictions were correct, and (2) that when the last response of a trial was an R+ it was repeated on 73.0% of the next trials, as opposed to 63.9% repetitions of R− choices. Voeks proceeded to perform three experiments as a more adequate test of the Guthrian hypothesis (now renamed *postremity*). For her first experiment (Exp. I) Voeks used a two-choice 18-unit Finger Maze of the type shown in Fig. 1. College students were run to mastery (three perfect trials) by the correction procedure. She reported 88.74% correct predictions via postremity. In Exp. II a two-choice 30-unit punchboard stylus maze and the noncorrection method were employed, resulting in 82.30% correct postremity predictions. This procedure was repeated in Exp. III with a threaded stylus, the postremity outcome being 83.05%. Since all three experiments involved two-choice tasks, the mean prediction index for 57 subjects of 84.69% is 34.69% better than chance. Voeks presented no separate analysis of frequency predictions, but she said that about 71% of her predictions were correct by postremity even when opposed by frequency, although this conflict reduced postremity's accuracy. Curiously, there was no analysis of the reinforcement factor. With respect to R+ vs. R− scores she asserted: "by the present hypothesis, postremity is neutral, having no differential effect on the correct and incorrect responses" (Voeks, 1948, p. 497).

A third investigation of this problem (Waters & Reitz, 1950) used a two-choice 15-unit Finger Maze on which 20 subjects were each given 15 trials by the noncorrection procedure. Half of the group received a changing stimulus situation from Trials 11 through 15 to see whether postremity predictions would be adversely affected. Waters and Reitz reported an average postremity score of 70% correct predictions based on 4200 responses. This figure, 20% over chance, is doubtless smaller than Voeks' margin because the subjects were not carried to mastery. A more interesting aspect of the Waters–Reitz experiment, however, is their trial-by-trial analysis. Since the constant and variable groups were not significantly different on the last five trials (an embarrassing result for the Guthrie-Voeks viewpoint), they may be combined. When I did this I found that the percentage of correct predictions increased with stage of practice, rising from 53% on Trial 2 to 80% on Trial 15. Extrapolating to the Peterson–Voeks criterion of three perfect trials in a row, it is obvious that postremity would eventually predict with 100% accuracy if practice were carried far enough. Like Voeks, and Peterson earlier, Waters and Reitz made no analysis of reinforcement predictions. Clearly, more extensive comparative data of a detailed nature are needed.[7]

---

[7] Overall and Brown have offered cogent arguments and facts in support of multi-

In an effort to contribute information toward this objective, I have reported two theoretical studies on the Selective Mathometer (Noble, 1962, 1965b) which evaluate the relative efficacy of the postremity, frequency, and reinforcement principles under correction, noncorrection, and modified-correction procedures. The first study (Noble, 1962) used data collected from 100 subjects performing on standard 10-choice 10-unit Mathometer problems for 20 trials (Noble, Alcock, & Noble, 1958b). Half the subjects practiced by the modified-correction method (permitting alternative corrective responses, but only for the 2-second duration of a slide); the remaining 50 practiced by the noncorrection method (requiring one and only one response per slide). Both groups were experimenter paced. From the graphic records of 20,000 choice points, individual pushbutton behavior was analyzed under the two $S_e$ conditions. Extending and improving the scoring method developed by Voeks, I made three independent predictions for each subject's first response (R) at each choice point beginning with Trial 2. The three predictions were based on the following specific hypotheses derived from classical association theory.

1. *Postremity:* the last nominal (numbered) R occurring at a given choice point (whether R+ or R–) on the immediately preceding trial is predicted, regardless of consequences (after-effects, feedback, etc.), a prediction being made at every choice point.

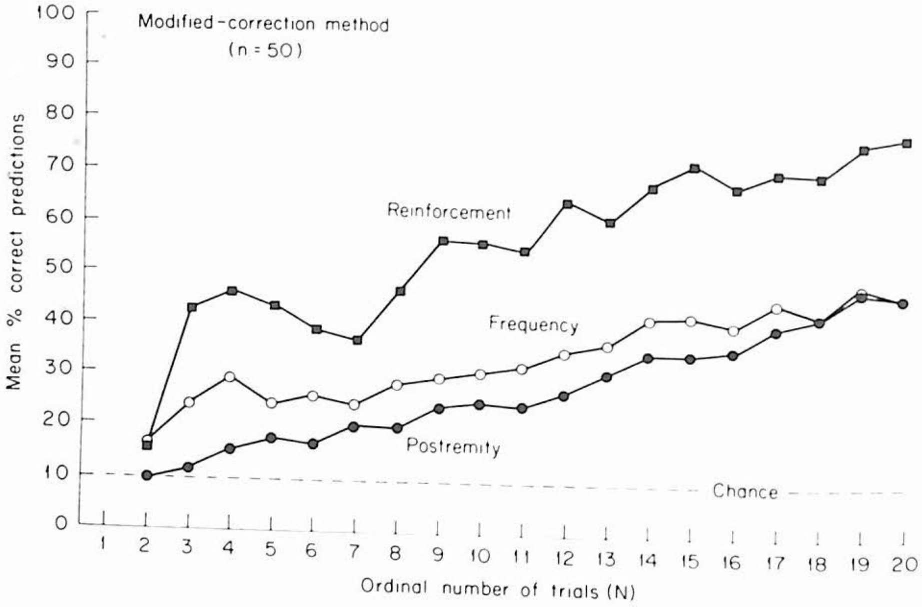
2. *Frequency:* the nominal R occurring most often at a given choice point (whether R+ or R–) on all preceding trials is predicted provided $R_p > .50$, no predictions being made if the frequencies are equal.

3. *Reinforcement:* the nominal R most frequently reinforced by the green light ($S_g$) at a given choice point on all preceding trials is predicted provided $R_p > .50$, no predictions being made if the frequencies of reinforcement are equal.

To be as fair as possible to the three contenders, a few ground rules were adopted in advance of scoring. For reinforcement predictions, late errors were counted but late (unreinforced) corrects were ignored; in frequency and postremity predictions, both late R+ and late R– choices were counted. All double Rs were scored as singles, but no intertrial Rs were scored. Omissions were counted as disagreements only if the hypothesis predicted an overt R.

---

factor theoretical models of the postremity-frequency type (1957) and the cognitive-postremity type (1960). Since their papers deal primarily with the behavior of a-mazed rats faced with Hamiltonian insoluble problems, I shall refrain from discussing them here. It is interesting, however, to note their contention that subjects "will follow the *hypothesis* which most *recently* led to the desired *goal*" (Overall & Brown, 1960, p. 191; italics added). This eclectic viewpoint represents a hybridization of Tolman (1932), Guthrie (1952), and Hull (1943).



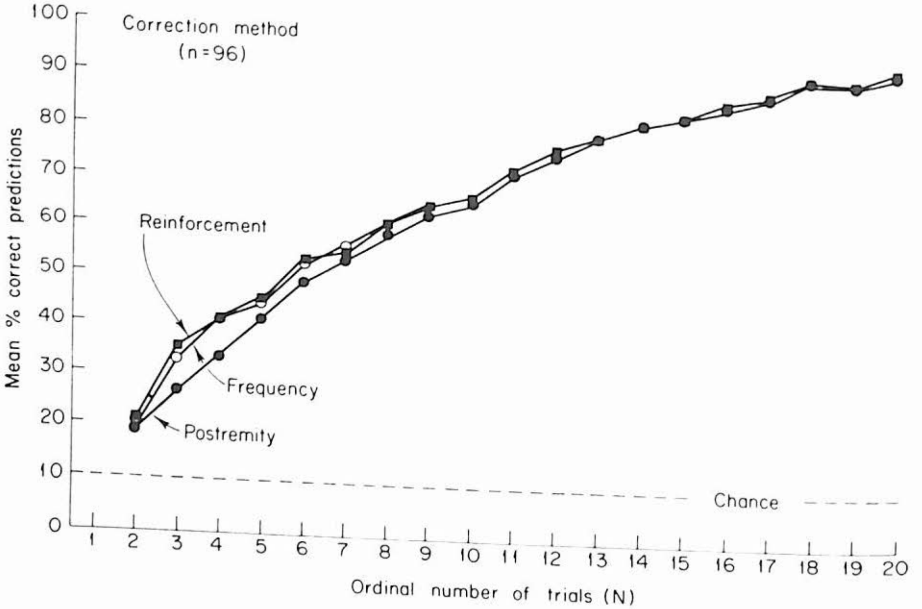
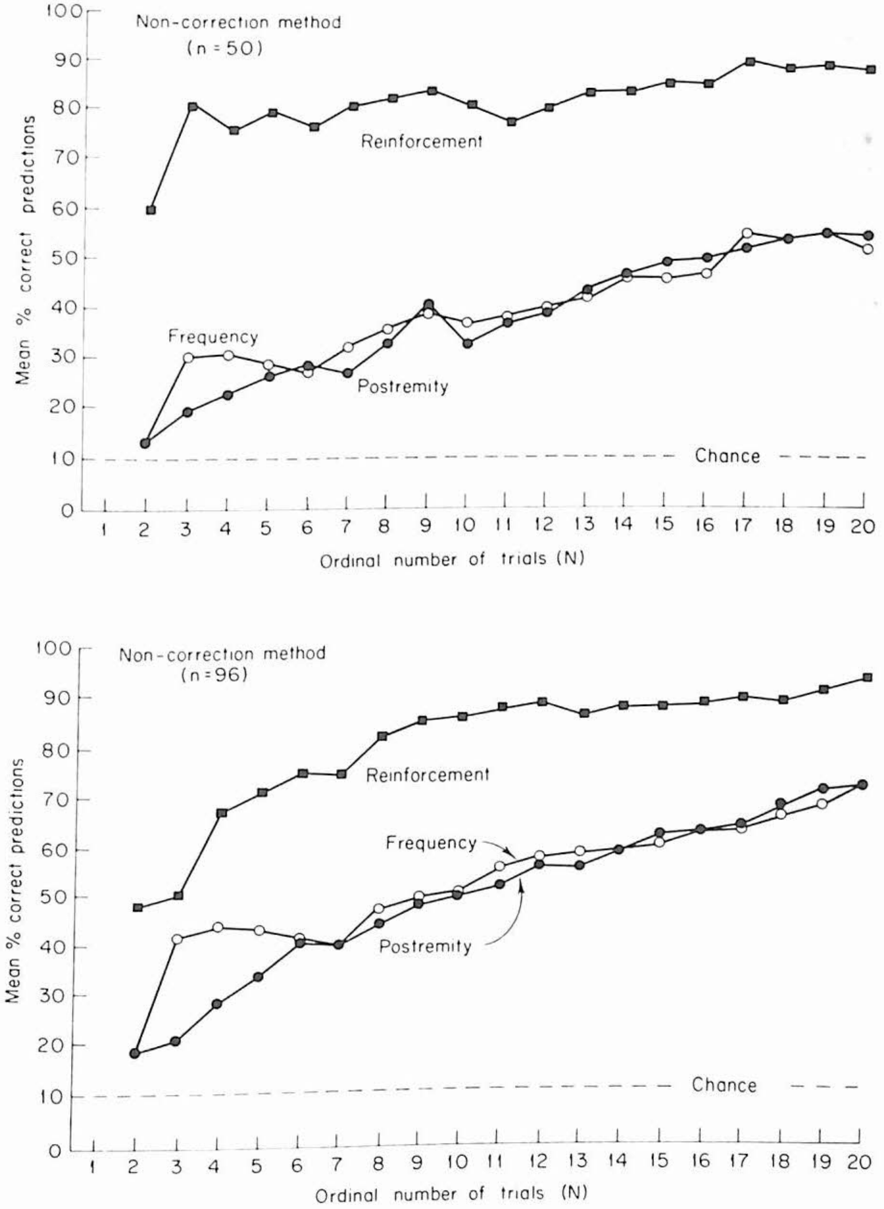


FIG. 4. Comparative predictabilities of the reinforcement, frequency, and post-



remity hypotheses under three methods of practice as functions of stage of practice (after Noble, 1962; Noble, 1965b).

The percentage of agreement between prediction and fact was then computed for each hypothesis by subjects and by trials, with the results shown in Fig. 4. Note that the ordinate scales display the performance not of subjects but of hypotheses. Looking first at the *modified-correction* method (upper left graph), we see that reinforcement predictions are generally far superior to predictions by frequency or postremity, the latter being definitely inferior to frequency. Looking next at the *noncorrection* method (upper right graph), an even greater superiority of the reinforcement hypothesis is evident. Although frequency is still ahead of postremity, its margin of superiority is less.

My second study (Noble, 1965b) used data collected from 192 subjects performing on the same kinds of 10-choice Mathometer problems as in the 1962 report. In this experiment (Noble, 1965a) half the subjects practiced for 20 trials by the correction method (requiring corrective responses until reinforcing feedback was received); the remaining 96 practiced by the noncorrection method as defined earlier. In other words, the correction group was subject paced whereas the noncorrection group was experimenter paced. From the graphic records of 38,400 choice points, individual responses were analyzed under the two $S_i$ conditions, as described above.

Returning now to Fig 4 and looking first at the *correction* method (lower left graph), we see that there is little difference in the predictive efficacy of the three hypotheses. Reinforcement still leads frequency (albeit by a tiny margin) and both are superior to postremity until the last quarter of practice when they merge above the 80% ordinate. Looking next at the *noncorrection* method (lower right graph), we find essentially a replication of the earlier investigation. In general, Fig. 4 shows that reinforcement predictions are consistently and decisively more accurate than all others, with frequency winning out over postremity in a close race. All three hypotheses improve in validity as practice continues, and would probably converge if mastery were approached.

The overall results of these four independent comparisons, pooling the observations of 292 subjects responding one or more times at 58,400 choice points,[8] are summarized in Table I as mean percentages of correct predictions during Trials 2 to 20. With the random accuracy level at 10%, the reinforcement principle predicts 70.98% better than chance when the two noncorrection groups are averaged. Frequency and postremity exceed chance by only 36.04% and 33.76%, respectively. We may conclude that in the noncorrection situation, reinforcement's accuracy for a 10-choice task is about

---

[8] Assisting in these analyses without benefit of automation, serving either as first computer or principal checker, were Ridgely W. Chambers, Joan E. Eldodt, James E. Fuchs, Helen R. Jones, Don E. Lawrence, and Ella M. Rowe.

TABLE I

MEAN PERCENTAGE CORRECT PREDICTIONS BY THREE HYPOTHESES UNDER THREE METHODS OF PRACTICE DURING TRIALS 2–20 ON THE SELECTIVE MATHOMETER[a]

| Method | Hypothesis | | | $n$ |
|---|---|---|---|---|
| | Reinforcement (%) | Frequency (%) | Postremity (%) | |
| Noncorrection | 81.74 | 39.89 | 38.08 | 50 |
| | 80.12 | 52.18 | 49.43 | 96 |
| Modified-correction | 54.82 | 34.04 | 27.81 | 50 |
| Correction | 67.29 | 66.79 | 65.17 | 96 |

[a] After Noble (1962, 1965b).

twice that of its rivals. After five trials we can predict around 75% of the time what the initial response at the next choice point will be merely by keeping a record of the subject's reinforcement history. By Trial 18 the predictability is close to 90%. Using frequency or postremity, our predictions on those same trials would average only 33% and 60%, respectively. For analytic purposes, in consequence, maximum differentiation among the three hypotheses is provided by the noncorrection procedure, followed by the modified-correction and correction procedures in that order. In the correction method, since a reinforced response must always be last, no one will be surprised to find postremity predicting almost as well as reinforcement. Similarly, since the reinforced response tends with practice to become the most frequent response in the correction procedure, frequency also is able to predict at about the same level as reinforcement. Regardless of practice method, however, the validity of the reinforcement hypothesis for selective learning is strikingly demonstrated in these analyses. At the same time they appear to present a formidable challenge to nonreinforcement theories (Guthrie, 1952; Harlow, 1959; Voeks, 1948) and to theories which construe learning as a random event (Bower, 1959; Bush & Mosteller, 1955; Estes, 1959; Restle, 1962).

With respect to the *contiguity-reinforcement issue*, it is instructive to observe that while all three hypotheses assume S-R contiguity, two of them are distinctive in exploiting R-R components (e.g., correlating the *last* or the *modal* response with the *next* response). The power of the S-R-contiguity–frequency-reinforcement hypothesis, to enumerate all its elements, is that it includes an extra experimental component: the operation of reinforcing $S\text{-}H \rightarrow R+$ connections. As Seward *et al.* (1944) found in a straightforward S-R-contiguity–recency pushbutton experiment of the type Guthrie's (1952) theory would propose, the paradigm S→$R_1$ followed by S→$R_2$ failed to produce $R_{p_2} > R_{p_1}$. Unlike Voeks' postremity, then, reinforcement is *not* neutral: it *does* exert differential influences on R+ and R− choices. This is what accounts for the massive increments in going from the frequency to the reinforcement curves in Fig. 4 (for all but the correction procedure). Additional factors, which are considered later in the chapter, are the significant effects of generalization and nonreinforcement. By combining all these elements, as we are attempting to do in the Georgia laboratory, I believe that the average predictability of individual behavior (for carefully selected initial and boundary conditions) can be extended well into the 90% range.

With respect to the *continuity-noncontinuity issue*, there are few relevant human-learning investigations apart from a spate of recent verbal experiments on the role of item selection and other artifacts in producing ostensible "one-trial-learning" results. As Kendler (1959) foresaw and Estes (1962) has recognized, the methodological defects of the early paired-associate substitution designs have been exposed (Underwood & Keppel, 1962). To the discomfiture of most S-S as well as some S-R noncontinuity theorists, those star-crossed studies turned out to be exercises in the control of performance factors rather than proofs that $H$ is saltatory. While no dogmatic assertion is possible yet on the question of all-or-none vs. incremental principles of habit formation, it is manifest that advocates of the noncontinuity viewpoint have not exactly improved their defensive posture *vis-à-vis* those of the continuity persuasion. The incrementalists are in a position of strength for several reasons: (1) most empirical group acquisition curves measured in terms of $R_p$, $R_a$, or $R_{1/t}$ can be described as continuous positive functions of $N$ (Hull, 1943; Woodworth & Schlosberg, 1954); (2) the results of animal discrimination studies have generally favored the continuity view (Deese, 1958; Kimble, 1961); (3) the rapidity of changes in learning-set scores among both primate and human subjects is a cumulative function of the amount of interproblem training (Harlow, 1959); (4) human overlearning-retention experiments are consistent across memory drums and finger mazes in showing positive correlations and a diminishing-returns principle (McGeoch & Irion, 1952); (5) the occasional appearance of discontinuous $R_p$ scores can be handled within continuity theory by attenuating the oscillation function, as in cases of increased experimental control (Spence, 1956).

For a sampler of human selective-learning experiments supporting the incremental principle, I shall mention three. Prentice (1949) adapted the classic two-choice cue-reversal discrimination-learning paradigm for college students, finding the reversal group's performance inferior to that of the nonreversal group. This is as it should be if erstwhile reinforced choices must be extinguished and former avoidance responses replaced by approach when the cues are altered for the reversal subjects. In a similar investigation Walk (1952) also got continuity results from a pair of discrimination-learning experiments, although the negative transfer effects were brief. Finally, there is a demonstration by Paul and Noble (1964) of the positive transfer effect of successive habit reversals on the Mathometer which turned up some evidence inconsistent with the ahistorical postulate of Gestalt theory. We noticed that the number of subjects attaining perfect scores on the second postreversal trial of each block increased steadily with amount of reversal training. These data were reanalyzed and suggest that although insight may be a within-subject phenomenon, it can be manipulated like any dependent variable, then assessed as a between-subject effect. In sharp refutation of the noncontinuity credo that insight is independent of repetitive training, the probability of one-trial reversals occurring is a continuous positive function of $N$.

To recapitulate, the odds-on favorite in the learning lottery is a reinforcement theory which assumes an underlying continuity for the growth of associations. Now, in the final section, I turn to a review of the laws of selective learning, prefacing the discussion of each group of phenomena by a tentative coordination with our approximation formula.

## Laws of Selective Learning

### COMPLEXITY OF TASK: $R = f(S_c)$

As a class these $S_c$ factors affect the parameters $i$ and $r$ of Eq. (4) plus the variance of $R_p$, and they are believed to be nonassociational performance variables.

Evidence supporting this hypothesis is supplied by several Mathometer experiments. In one (Noble, 1955) the number of available responses ($N_R$) was varied between 4 and 10 in 336 subjects to generate a set of seven four-unit problems (cf. Fig. 3). There were significant differences in trials to mastery among the groups. Subsequent theoretical analysis (Noble, 1957a) revealed that the $R_p$ curves were skewed and S shaped; the origins agreed with rational expectation ($i = 1/N_R$); the inflection points on the $N$ scale varied directly with the number of exposed keys; and the acquisition rates varied inversely. The average goodness of fit using Eq. (7) to secure $F$ values was 99.1%. An extension of this investigation (Noble, 1957b) on four new groups of subjects, exploring the effects of $N_R$ from 11 to 14, also produced significant results. Earlier (Noble, 1954) I had recommended that *complexity* be regarded as the total possible number of different solutions to a problem and defined by the algebraic formula for permutations ($nPm = n!/[n-m]!$). Measuring *difficulty* by the median trials needed to reach different criteria of mastery led to a logarithmic difficulty-complexity relationship, with $nPm$ ranging from 24 to 5040 and a mean $F$ value for the four logarithmic equations of 96.5%. Kaess and Zeaman (1960) obtained significantly greater errors in a punchboard task when response availability was increased from $N_R = 1$ to $N_R = 5$. Then there are tasks which are slightly beyond our taxonomic borders which produce similar results. Riley (1952) for paired-associate verbal learning and Neimark (1956) for Humphreys-type verbal expectations both found decrements when they increased the number of choices.

Another dimension of task complexity is encountered when the number of choice points $(N_P)$ is covaried with $N_R$. Brogden and Schmidt (1954a,b) used the correction method with verbal mazes varying from 2 to 12 choices and from 16 to 24 units. Difficulty, measured by time and errors to criterion, increased proportionately with $N_R$ and there was an interaction between $N_P$ and $N_R$. A model was developed for these data by Simon (1957) which accounted for 93.7% of the variance in error scores, but it predicted independence of trials-to-criterion scores from $N_R$. Meanwhile I had engaged the Complex Coordinator (cf. Fig. 1) in a related enterprise covering a range from 5 to 40 matches, emerging with a rectilinear time-length equation whose $F = 99.9\%$ (Noble, 1953a). These linear laws, I suspect, are special cases of task and method. Perhaps they obey a power function in the general case, with exponents of unity applicable here.

As a further test of Eq. (4) I used the modified-correction method on the Mathometer, jointly varying $N_R$ and $N_P$ from 4 to 14 (Noble, 1957c). There were no serial position error gradients. The $R_p$ scores behaved as expected, with $i$ and $r$ accounting for both response availability and task length, giving a mean $F$ of 98.8%. Since the form of Eq. (4) was invariant over changes in task length, the $R_p$ data agreed with McGeoch's version of the Kjerstad–Robinson hypothesis (Noble, 1957c, p. 251). Equally critical for theory, however, was the form of the length-difficulty relationship for trial durations. It was parabolic, supporting Thurstone's (1930) hypothesis with $F = 99.8\%$. Median total time to mastery $(R_{\dagger})$ won out over several other measures of difficulty. By making some simplifying assumptions in Thurstone's formula, I reduced it to the form

$$R_{\dagger} = t(N_P)^b \tag{8}$$

where $t$ = trial duration in seconds; $b$ = slope parameter $> 0$. Another Mathometer experiment (Noble & Fuchs, 1961) confirmed and extended these findings upon 400 subjects in 20 combinations of $N_P$ and $N_R$. Difficulty increased with either variable, and the effects of practice, conditions, and their interaction were all significant. By making certain keys correct at more than one choice point, we forced the subjects to alternate their choices. Alternation joined forces with irrelevancy to produce maximum difficulty.

Uniformity of the S-R sequence in mathometry was investigated by Farese and Noble (1960), who predicted that a constantly changing order in a four-choice four-unit task would be more difficult than the standard one. With only $r$ allowed to vary, Eq. (4) gave an average $F$ of 97.96%. In addition to decremental effects on means there were inflationary effects of $S_c$ on variances. Our error curves are reproduced in Fig. 5 to show the greater intertrial fluctuations of the variant group. The residual $\sigma^2$ of this group was about three times that of the invariant group, and early trials showed greater variability than later trials. This outcome is consistent with the Hull–Spence concept of $I_{o}$, especially since the early-late difference in variability was also greater for the variant group. A related study of sequential complexity was done by Barch (1960a). He employed his Serial Reactor (Barch, 1960b) to find out whether the $S_c$ factor was significant. It was, and it also interacted with practice. Color-form cue irrelevancy was not significant. Loess and Duncan (1952) studied two-choice discrimination learning under simultaneous and succes-

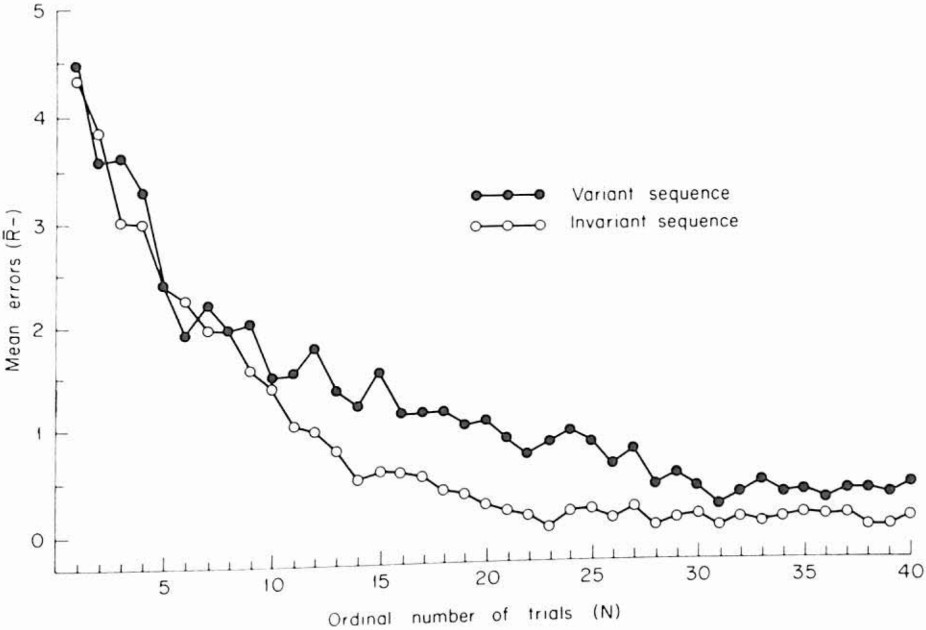


FIG. 5. Mean error curves (R̄−) in selective learning as functions of the number of practice trials ($N$), with uniformity of the S-R sequence ($S_c$) as the parameter. Each curve is based on the data of 48 subjects (after Farese & Noble, 1960, p. 118).

sive presentations. Although there was no difference for easy problems, the simultaneous condition was superior for difficult problems. The implication that subjects perform better on close choices when they have an opportunity to compare S+ and S− was confirmed on children (Perkins, Banks, & Calvin, 1954). Applying Spence's (1960) notion of successive cue-position patterning to account for simultaneous discrimination learning, there are experiments on rats (Birch & Vandenberg, 1955) and children (White & Spiker, 1960) which show that responses to patterns occur under conditions of differential reinforcement of the component stimuli. White and Spiker noted that amount of patterning is significantly affected by increasing the similarity among the nonspatial cues within a trial.

That facet of $S_c$ variables concerned with display-control relationships in perceptual-motor tasks is a favorite among engineering psychologists. Lest this be regarded as alien to the field of selective learning, we have Fitts' (1964) conclusion that skill is composed of fundamentally discrete components. His review suggested that the study of separate S-R elements will make important contributions to the theory of continuous action, including spatio-temporal patterning, motor coordination, serial communication, verbal behavior, thinking, and decision processes. On a variety of devices Fitts and his colleagues have shown that direct S-R compatibility is most efficient. Independent data on this phenomenon were supplied by Adams (1954c). Positive transfer between tasks was initially related to the degree of similarity in the display-control relationships. Morin and Grant (1955) confirmed a number of earlier results on the Multiple Serial Discrimeter (cf. Fig. 1). Using rank correlation (tau) as a measure of S-R correspondence, they obtained clear facilitation in transferring subjects from a noncorresponding task to one with high S-R compatibility. Nystrom and Grant (1955) assessed the compatibility of five angular orientations of lights and keys, finding the greater the angular displacement from zero the less the compatibility and the poorer the performance, especially beyond 90°. In general, the magnitude of S-R compatibility effects varies positively with task complexity (Fitts, 1964). A monograph by Seibel (1962) described an investigation of discrimination reaction latencies under variations in $S_c$ defined by the number of S-R pairs. A five-light five-key one-hand apparatus provided $2^5 - 1 = 31$ S-R alternatives. Contrary to information theorists' expectations, $R_t$ scores were not a positive linear function of information transmitted in bits (cf. also Adams, 1964). Instead, the obtained function was negatively accelerated, reaching an asymptote around 3 bits (8 alternatives) and holding steady up to 5 bits (31 alternatives). A subsequent 10-light 10-key study (Seibel, 1963) employed both hands to explore $2^{10} - 1 = 1023$ S-R alternatives during 200 experimental sessions. As expected from his first study, Seibel found the mean $R_t$ of the maximum complexity condition only slightly different (< 25 milliseconds higher) than that of the 31-alternative condition.

My last examples of $R = f(S_c)$ functions involve comparisons of the correction and noncorrection methods. They have interesting implications for reinforcement theory and for programmed-learning technology. Hull and Spence (1938) verified a prediction that the correction procedure would produce faster trial-and-error acquisition in rats. The reason: the subjects received both reinforcement of R+ and nonreinforcement of R− on each trial, thus accelerating the learning process. Irion and Briggs (1957) trained 240 high school students on 20 S-R combinations under serial and paired-associate presentation using the Subject-Matter Trainer (Besnard *et al.*, 1955). There were significant differences favoring a simple correction procedure ("practice mode") over a noncorrection procedure ("single-try mode"), and two other complex correction methods ("quiz mode," "modified-quiz mode") were even more effective. Then there is a recent Mathometer experiment (Noble, 1965a) in which 192 matched college subjects practiced a 10-choice 10-unit problem under one of the two methods for 20 trials. I found significant effects of practice, method, and their interaction. When $R_p$ scores were measured the correction method was superior over all trials; the two curves fitted by Eq. (4), varying only the $r$ parameter, gave an average $F = 99.50\%$. When mean R− scores were measured the noncorrection (apparatus-paced) subjects produced significantly fewer errors until the last five trials. At that point the fast-falling curve of the correction (self-paced) subjects began to cross under the noncorrection curve. There is no inconsistency. With more responses allowed the correction group committed more errors than the noncorrection group, but they were eliminated at a faster rate. Clarification comes by plotting error percentages; when this check is made the correction group's curve is superior throughout, in exact agreement with $R_p$ scores.

### FEEDBACK CONTINGENCIES: $R = f(S_g)$

As a class these $S_g$ factors affect the parameters $r$ and $a$ of Eq. (4) plus the extinction of $R_p$, and they are believed to be both learning and performance variables.

Experimental data relevant to rate of acquisition $(r)$ appeared in a Mathometer study on the influence of different loci of reinforcing feedback in 96 military subjects (Noble & Noble, 1958). One group (serial $S_g$) received immediate reinforcement after each R+ whereas the other group (terminal $S_g$) had to wait until the end of the trial. The task was a four-choice four-unit problem practiced by the noncorrection method to mastery. We found decisive superiority for the serial reinforcement condition. With $a$ and $i$ held constant while $r$ alone was allowed to vary, Eq. (4) predicted well, the mean

*F* was 99.30%. Despite the fact that $R_l$ data increased with training, latency scores were not significantly different between the serial and terminal groups; nevertheless, the intraserial $R_l$ scores of the terminal group declined systematically toward the end of each trial (the point of reinforcement). Again, we saw no serial position effects for R+ scores. A sequel to this was an experiment on delayed serial feedback (Noble & Alcock, 1958). We recruited 576 subjects and tested them on two different problems, dividing them into six $S_g$ delay groups in the range from 0 to 3 seconds. The problems differed in complexity, one being a four-choice four-unit task, the other a 10-choice 10-unit task; all subjects practiced for 20 trials by the noncorrection method. Differences in acquisition speed due to $S_c$ were significant and so was the $S_c \times S_n$ interaction. There was no appreciable variance attributable to $S_g$ delay or to practice interaction in the short task. The long task gave some hints of increasing impairment with greater delays (including a significant $S_n \times S_g$ interaction which affected the *r* parameter), but the main effect of $S_g$ was slight. For both tasks we noted that mean $R_p$ scores during Trials 1–10 rose slightly from 0 to 1 second, then fell from 0 to 3 seconds (Noble & Alcock, 1958, Table 2). In another Mathometer investigation (Fuchs, Leland, & Noble,1960) 230 college students practiced for 90 trials by the noncorrection method on a four-choice four-unit task. They were given five different terminal delays from 0 to 16 seconds, with and without proportional increases in the number of interpolated responses (left-to-right button pushing) between trials. There were no decrements. The contention that human subjects are highly resistant to interference from delayed feedback (Bilodeau & Bilodeau, 1961) is not contested in this laboratory. Merely delaying $S_g$ seems to be irrelevant unless effective response competition is produced (e.g., by the terminal reward procedure). When interference occurs the *r* parameter is suitably affected.

Many varieties of $S_g$ operations are reinforcing. Birch (1964) developed a puzzle task to measure instrumental approach speed in a two-choice situation whose alternatives differed in difficulty and/or incentive value. The $R_{1/t}$ scores were a positive function of trials, and there was an interaction between incentive value (amount of success) and spatial position. In the Morin–Grant (1955) experiment mentioned earlier, subjects given noncorresponding S-R matches on the Multiple Serial Discrimeter relied more heavily on informative feedback than those given corresponding S-R matches. Primary rewards have their place, too. Studying two-choice discrimination learning and transposition in children, Terrell and Kennedy (1957) found candy reward more effective than praise or reproof. There are two 10-choice Mathometer experiments on types and combinations of feedback. The first (Noble *et al.*, 1958b) was a 100-subject comparison of our standard serial reward condition with a condition in which we reported R+ scores to the subjects in addition to giving them green lights. The light-score group performed slightly but not significantly better than the light-alone group. In the second experiment (Noble, 1965d) 80 subjects were divided into four groups, all receiving the green light serially for R+. One group saw distinctive labels above the keys, another received verbal reports of the key's identity whenever R+ occurred, a third had both treatments, and the fourth was a control. The light-alone group came out best, the light-label-report group second, the light-report group third, the light-label group last. Analysis-of-variance tests revealed significant differences due to practice, feedback treatments, and their interaction. Apparently our standard $S_g$ conditions are optimal, for multiple feedback tends to be disrupting. Jones (1945) investigated trial-and-error learning in 144 junior high school students with a 40-hole punchboard maze. Different kinds of feedback signals (buzzer, light, tone, vibrator) produced different affective qualities for R+ and R– choices, yet had no significant influence on time or error scores. Mere knowledge of R– was as effective as feedback plus stimulus emphasis.

Nonreinforcement, that powerful alter ego of the law of effect, is an $S_g$ factor which tends to be somewhat neglected in theoretical psychology. Impressed by the Hull–Spence (1938) demonstration that rats can master a T maze quickest when reinforcement of R+ and nonreinforcement of R– are working together (correction procedure), Denny and Dunham (1951) contrived a neat proof of the hypothesis that four nonreinforcements per day for rats are better than one when the number of reinforcements is held constant at two. Nor is the principle limited to rodents, as G. N. Cantor and Spiker (1954) have clearly shown with preschoolers learning a discrimination task. The technique of forcing extinction trials for R– upon the subject definitely aids in his acquisition of R+.

From nonreinforcement of R– to partial reinforcement of R+ is a short step. Ninety-two collegians were assigned to the Mathometer by Taylor and Noble (1962) for a comparison of continuous reinforcement (100 percent $S_g$) with three partial reinforcement schedules (25, 50, 75 percent $S_g$). All subjects practiced a four-choice four-unit noncorrection task for 15 reinforced trials, then received 20 extinction trials. When we analyzed the acquisition data by successive reinforced responses, orderly decremental effects due to nonreinforcement appeared. Consistent with the augmented S-R reinforcement theory discussed in the section on Theory of Selective Learning, amount of impairment was perfectly correlated with the number of extra nonreinforced trials given. During extinction the continuous group dropped below the others initially, then at Trial 5 began a comeback to the level of the partial groups. This is a rare phenomenon in the partial reinforcement literature, doubtless because most experiments are terminated prematurely. (Small wonder; extinction is onerous for subject and experimenter alike.) There were no systematic differences among our partial groups in extinction when relative percentages of R+ were plotted against nonreinforced trials. Hubbard (1951), investigating secondary reinforcement in a three-choice discrimination task with 81 subjects, had also gotten nonmonotonic extinction functions. His curves fell precipitously at first, but after Trials 3–5 they rose until Trial 20 where they leveled off just like ours. In order to replicate the Taylor–Noble findings and to amplify this extinction phenomenon, I recently ran 184 subjects on the same four-choice task for 20 reinforced trials under 50 and 100 percent $S_g$ followed by 40 extinction trials (Noble, 1965c). In acquisition the 100 percent group was markedly superior to the 50 percent group, confirming the inhibitory effect of nonreinforcement and its correlated influence upon the *r* parameter. The extinction curves reproduced in Fig. 6 show that the 100 percent group reached its low point during Trials 3–4, after which it moved up toward convergence with the 50 percent group. By Trial 23 the two groups were statistically indistinguishable; about 70 percent of the subjects' responses, relative to their last two acquisition trials, were correct from then on. The concordance between the data of Taylor and

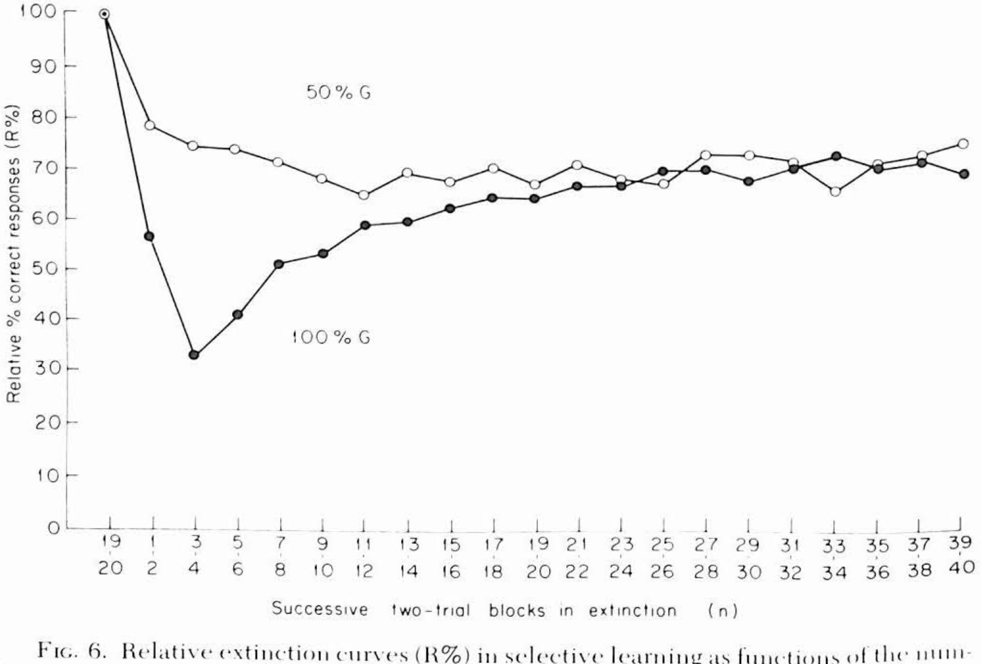


FIG. 6. Relative extinction curves (R%) in selective learning as functions of the number of nonreinforced trials (*n*), with original reinforcement schedule ($S_g$) as the parameter. Each curve is based on the data of 96 subjects (after Noble, 1965c).

Noble (1962), Hubbard (1951), and Noble (1965c) is reassuring, although the implications of the subjects' "strategies" are not yet clear. Related phenomena from a lever-positioning task have been observed by E. A. Bilodeau (1953); they are discussed in the context of shifting $S_g$ by I. McD. Bilodeau (Chapter 6).

For a final illustration of $S_g$ factors, consider their influence upon the asymptote parameter of Eq. (4). The quantity $a$ is the limit of $R_p$ and reflects the maximum proportion of reinforceable responses in a given task. By manipulating this proportion in a 10-choice selective learning situation from 1.00 down to .40, it should be possible not only to lower the asymptotic "ceilings" but also to depress the initial proficiency levels of groups subjected to different values of $a < 1.00$. Experimentally this can be easily achieved on the Mathometer by arranging to have one or more keys nonreinforceable (nonfunctioning). The theoretical reason for the expectation of lower initial probabilities for insoluble problems is that $a$ and $i$ are multiplicative. Thus, $R_p$ at the outset of training (Trial 0) is predicted by the *joint* probability that the subject will make a correct choice (R+) *and* be reinforced ($S_g$). Moreover, each inflection point should occur at an ordinate value represented by the ratio $a/e$, where $e = 2.718$ and $a$ varies from 1.00 to .40. Successful empirical verification was reported by Chambers and Noble (1961) with 150 subjects practicing a 10-choice 10-unit task for 50 trials. The $R_p$ curves shown in Fig. 7 are generally consistent with our pre-experimental predictions regarding $a$ and $i$. The $r$ parameters were obtained in the usual way after appropriate modifications of the graphical operations leading to Eq. (6). Predictability by Eq. (4) was high; the mean value of $F$ for the five curves was 99.30%. As an independent check on the influence of different "degrees of solubility" on selective-learning performance, Chambers and I computed a measure of difficulty (Noble, 1957a) which would be comparable among all groups. It was the median number of trials needed to reach a mastery criterion of four correct responses (five perfect trials in succession). Difficulty emerged as an inverse linear function of the degree of solubility (= 100$a$). To generalize, the greater the percentage of reinforceable responses in a series the easier it is to master a fixed number of these responses. By invoking the multiplication theorem of the probability calculus in this experiment, we were able to link Eq. (4) to the relative frequency theory of probability in a rational manner.

### WORK DISTRIBUTION: $R = f(S_u)$

As a class these $S_u$ factors affect only the $r$ parameter of Eq. (4), and they seem to influence learning by way of performance.

Most investigations of $S_u$ variables have been performed with motor-skill devices using $R_t$ or $R_a$ (amplitude) scores (Bilodeau & Bilodeau, 1961; Fitts,

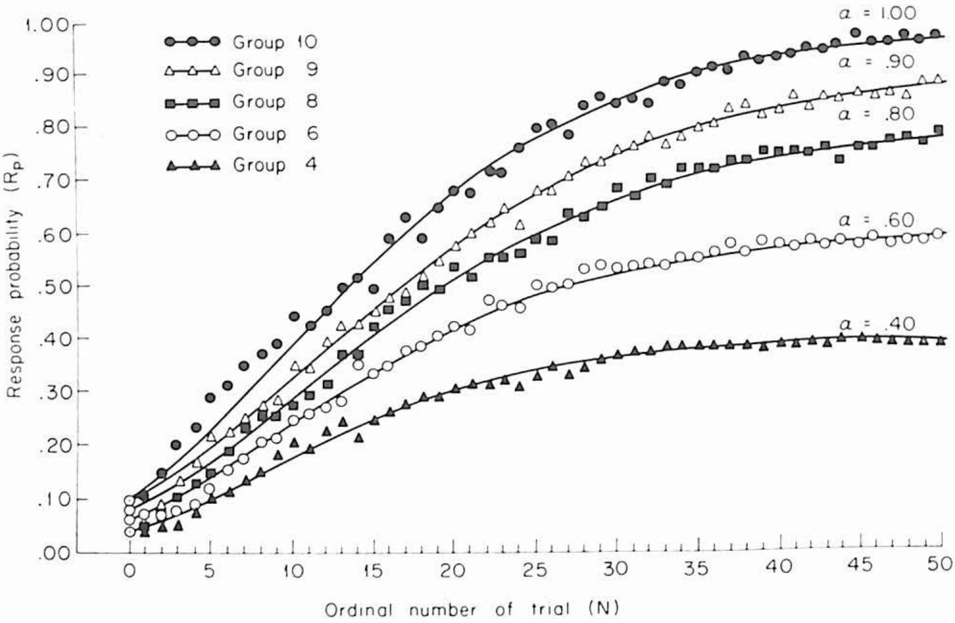


FIG. 7. Response probability *($R_p$)* curves in selective learning as a function of the number of practice trials *($N$)*, with different asymptotes *($a$)* as the parameter. For computations of the rational origins, see text. Each curve is based on the data of 30 subjects (after Chambers & Noble, 1961, p. 419).

1964). Such measures usually reveal marked initial and final differences in proficiency. Their disadvantage is their lack of analyticity. By contrast, the advantage of $R_p$ scores on the Mathometer is that comparative studies of work distribution among different learning groups can be carried out within a single coordinate system, as in the case of the $S_c$ factors discussed earlier. There is increasing agreement among quantitatively-oriented learning theorists that $R_p$ is the most analytic operational indicator we have of the relative strengths of response tendencies. Thus Restle wrote, ". . . each response may be assigned a 'strength,' and . . . the probability of a certain response is its relative strength" (1959, p. 419). Skinner (1961) and Estes (1962) have both argued convincingly for an emphasis upon $R_p$ measures. For these reasons, and because of the probabilistic nature of selective learning (cf. Fig. 2), I shall continue to mention Mathometer experiments prominently.[9] Studies of correction procedures (Noble *et al.*, 1958b; Noble, 1965a) and distributed practice (Noble & Taylor, 1959) have addressed themselves to some of the major performance variables.

---

[9] Prof. Noble's review of $S_u$ factors here was deleted by the Editor to save space. It will be included in his forthcoming *Annual Review* chapter, "The Learning of Motor Skills" (Noble, *Annu. Rev. Psychol.*, 1967, vol. 18).

TRANSFER OF TRAINING: $R = f(S_t)$

As a class these $S_t$ factors affect the parameters *i* and *r* of Eq. (4); depending on conditions they may produce either learning or performance changes.

We first studied transfer on the Mathometer, as a joint function of specificity of instructions and habit reversal (Noble, Alcock, & Farese, 1958a). A four-choice four-unit task was practiced by 48 subjects under two sets of instructions for 20 trials, then reversed for another 20 trials. The nonspecific instructions produced lower initial responding and a slower rate of acquisition as compared with our standard specific instructions. Following habit reversal both groups fell to a chance level, after which they showed a significant learning-to-learn facilitation in acquiring proficiency on the reversed series. The instructions' influence appeared to transfer positively to the postreversal period, and the specific and nonspecific groups were both approaching a common asymptote of unit probability at the end of the practice period. Then a more extensive study was conducted (Paul & Noble, 1964) upon 88 subjects practicing a 10-choice four-unit problem ($N_R = 10$, $N_P = 4$) under either a standard sequence for 80 trials or a constantly reversing sequence for eight blocks of 10 trials each. After transferring the subjects to a unique problem with four different keys on Trial 81, we found that the reversal group performed significantly better than the standard group. Both groups were given 10 trials on this new problem. There was a significant $S_n \times S_c$ interaction during the transfer period. These results confirm Harlow's (1959) predictions and are consistent with experiments by Adams (1954b), Duncan (1958), and Morrisett and Hovland (1959) which had suggested that a critical factor in determining the relative merits of learning-to-learn and learning sets for transfer of training is the effect of the number of *intra*task trials which *multiple*-problem trainees receive prior to the transition. Specifically, their proficiency levels should be comparable with those of single-problem trainees. The Paul–Noble study thus proves that maximum positive transfer to new problems occurs when optimal amounts of general as well as special training are given. Data of both standard and reversal subjects were fitted satisfactorily by Eq. (4), the mean *F* value for all 11 curves being 97.22%.

Lewis and his associates at Iowa performed a remarkable series of projects in the late 1940s and early 1950s with the Complex Coordinator (cf. Fig. 1), all addressed to the question of facilitation and interference. Mashburn's apparatus was modified to provide reversible control-display relationships plus independent R– and R+ counters. The modal design was ABA. As Bilodeau and Bilodeau quaintly styled this all-out quest for retroactive interference phenomena, "Lewis found only shards of the holy grail" (1961, p. 262). Lewis, McAllister, and Adams (1951) reported that practice both facilitated and interfered with the interpolated learning of reversed controls, while amount of retroactive interference depended upon the amount of original practice. A related study by McAllister and Lewis (1951) revealed that a similar effect could be produced by interpolated practice. The subjects exhibited a paradoxical increase in both R+ and R− scores when S-R relationships were reversed, and amount of interference was negatively correlated with degree of reversed practice for strong original habits.

A number of $S_t$ experiments have employed the Star Discrimeter (cf. Fig. 1). Here are some of the best. Duncan (1953) demonstrated how positive transfer increases with first-task practice and intertask similarity. That transfer was attributable to learning-to-learn and response generalization was established later (Duncan, 1955, 1958) in an analysis of bidirectional spatial error gradients around R+. The same apparatus was used by McAllister (1953) for an investigation of the influence of verbal paired-associate pretraining on nonverbal selective learning in 98 subjects. She found that relevant S-R pretraining facilitated later performance.[10] Later, J. H. Cantor (1955) got significant evidence for the superiority of relevant over irrelevant pretraining, together with a suggestion of decreased stimulus generalization. The Star Discrimeter, incidentally, produces a clearcut inflection point when $R_f$ scores are plotted against single trials; thus McAllister's acquisition curves resemble Mathometer data more closely than do those of Duncan or Cantor, who resorted to coarser grouping along the abscissa. McFann (1953) studied proactive and retroactive facilitation and interference. He found both positive and negative influences when practice on one task was followed by practice on a different task having elements in common. Pursuing Duncan's work further, Montgomery (1953) essayed a distributed-practice transfer experiment with null results; there was no differential transfer due to $S_u$ operations applied to first or second task. Retention of transfer effects after one day vs. 14 months away from the Star Discrimeter was investigated by Duncan and Underwood (1953). Positive transfer covaried with degree of practice on Task 1 and with intertask similarity, confirming Duncan's (1953) data. There was a little forgetting of Task 2 over one day, followed by proactive facilitation in relearning. Greater forgetting occurred after 14 months, but 60% of the amount forgotten was regained in 20 trials. The effects of similarity decreased with successive periods of practice and rest. Finally, Lewis and Miles (1956) returned to the stage with an ABA design for the Star Discrimeter using different light-slot combinations for Tasks A and B. In six groups of female subjects, all of whom received equivalent amounts of practice on Task A, five groups were given different amounts of practice on Task B, then relearned Task A. The experimental subjects suffered a bit of retroactive interference during early relearning trials, the amount varying with interpolated practice. In general, the Complex Coordinator and Star Discrimeter data are compatible. By minimizing the degree of manipulative skill required of their subjects, Lewis and Miles indirectly garnered data relevant to the Complex Coordinator as a selective learning device (cf. p. 55 above).

---

[10] The positive transfer value of verbal instructions prior to training on the Discrimination Reaction Timer was established by Noble, Alcock, and Frye (1959a). Since performance on this apparatus is known to be correlated with that on the Selective Mathometer (Noble *et al.*, 1958c), further evidence is adduced for my classification scheme in Fig. 1 and on p. 55 above.

A residue of $S_f$ experiments employing other apparatus remains. Spiker and Holton (1958) developed a six-lamp, nine-key motor paired-associate device to study interference in A-B, A-C designs. From the A-C condition they obtained significant interference which was an increasing function of the amount of first-task practice. Mandler (1954) had subjects press a sequence of six levers in a hexagonal arrangement in response to eight letters as stimuli. Different proficiency criteria (0, 10, 30, 50, or 100 error-free trials) were required for the training conditions, whereas Brucean transfer tasks plus a warm-up control made up the second phase. Mandler's results, bolstering the hoary Wylie hypothesis, were these: (1) learning to respond with a new R to an old S produced initial negative, later zero transfer; (2) making an old R to a new S conferred positive transfer, which increased with degree of original training; (3) making an old R to an old S caused no difference for low levels of original training, but facilitation increased with degree of overlearning. Recently, Mandler and Kuhlman (1961) fabricated a more complicated gadget with 64 switches in an 8 × 8 array, although only eight were used and the stimulus letters dispensed with. Five groups of subjects received different amounts of training on two different sequences which were also relearned. Positive transfer occurred from one sequence to the other, and the overlearned tasks were more easily relearned than those merely mastered. Overlearning of one sequence increased the retroactive interference with relearning of the other but it facilitated performance in relearning the same sequence. Zeiler (1964a) compared "component" vs. "configurational" learning in children. He discovered that 3-year-olds learned a two-situation discrimination problem primarily on a configurational basis (the same directional R was correct in training as in testing), while 5- and 7-year-olds learned on a component basis (the same stimuli were correct in the two phases). The nonreversal shift from components to configurations was harder for the older subjects. In another discrimination learning $S_f$ study (Zeiler, 1964b), 240 adult females received simultaneous vs. successive stimulus presentation on an intermediate size problem, followed by tests varying over five distance steps from the training stimuli. Successive presentation produced absolute transfer effects, as Spence's theory predicts, whereas the simultaneous conditions gave data ranging from predominantly absolute choices to relational choices as a function of increasing training-test differences. Why there was a return to "transposition" behavior following the overlap of the absolutistic and relational curves was not explained. Extending the cue-disparity techniques developed on primates, Murphy and Miller (1959) entertained school children with a two-choice discrimination problem in which stimulus (S), response (R), and reward (G) were either contiguous or separated. Cue separation turned out to be a significant source of impairment. This was recently corroborated by Jeffrey and Cohen (1964) using five combinations of S, R, and G with two age levels (41 vs. 53 months). Once more the old refrain: S-R contiguity is a necessary (though not sufficient) condition for learning.

Another dimension to $S_f$ variables is spatial generalization. Brown, Bilodeau, and

Baron (1951) examined this factor by recording voluntary Rs to visual stimuli (seven lamps in a row) which varied in their physical loci. The subjects were instructed to react only to the middle lamp (forsaking all others) by executing a simple finger-lifting R. Plots of $R_p$ as a function of visual separation in degrees for the trained (center) and test (peripheral) lights yielded smooth symmetrical generalization gradients. Omitting the inhibitory instructions and resorting to angular movements, Andreas (1954) got comparable results; then came two reports by Marx (1957a,b) containing enough spread-of-effect R− data to tickle the fancies of all devout Thorndikeans. A newer development in Brown's laboratory (Brown, Clarke, & Stein, 1958) is the horserace betting game, which generates nice summated generalization curves over the middle stimulus, as Hull's (1943) theory specified and as other experiments attest (e.g., Bilodeau, Brown, & Meryman, 1956; Kimble, 1961, p. 349). The Selective Mathometer, likewise, is a source of spatial generalization data. Relative frequencies of R− choices to the left and right of R+ were computed from an experiment (Noble *et al.*, 1958b) in which 50 subjects practiced a 10-choice task by the noncorrection method for 20 trials. Similar gradients have been seen when the modified-correction method is employed.

## Individual Differences: $R = f(O_h, O_s, O_y, O_d)$

As a class these organismic or *O*-factors [habit aptitude $(O_h)$, sex $(O_s)$, age in years $(O_y)$, individual drive $(O_d)$] affect the parameters $a$ and $r$ of Eq. (4), and they are believed to be performance variables.

Our first individual-difference study on the Selective Mathometer (Noble *et al.*, 1958c) was performed on 261 college subjects to determine whether selective learning proficiency is predictable from extra-task measures of aptitude $(O_h)$. A battery of eight printed and three apparatus tests uncovered multiple correlations ranging from .32 to .62, depending on the factorial nature of the tests and the complexity of the two 10-choice criterion Mathometer problems (a 4-unit and a 10-unit task). All scores were transformed to normalized *T*s. Consonant with earlier work by Woodrow (1938), Adams (1953), and others (cf. Bilodeau & Bilodeau, 1961; Adams, 1964), adjacent-trial correlations were higher than remote ones and external tests predicted advanced proficiency better than they did initial proficiency. Validity was highest for R+, lowest for R− scores. The best predictors were discriminative, spatial, and reasoning tests, notably the Discrimination Reaction Timer (cf. Fig. 1). There were no significant differences due to sex $(O_s)$. With respect to complexity, the apparent effect of lengthening the task (from $N_p = 4$ to 10) was to increase the factorial diversity of the criterion, although this interpretation is moot (cf. chapters by Irion and Jones). Figure 8 is nonetheless consistent with Hull's (1945) hypothesis in that $O_h$ influences the constants rather than the forms of behavior equations. By stratifying into four aptitude groups the 106 subjects who practiced the more complex task, then applying Eq. (4) with $a = 1.00$, $i = .10$, and only $r$ varying, we obtained a mean $F = 97.1\%$. Ability and variability varied inversely. A recent experiment (Noble,

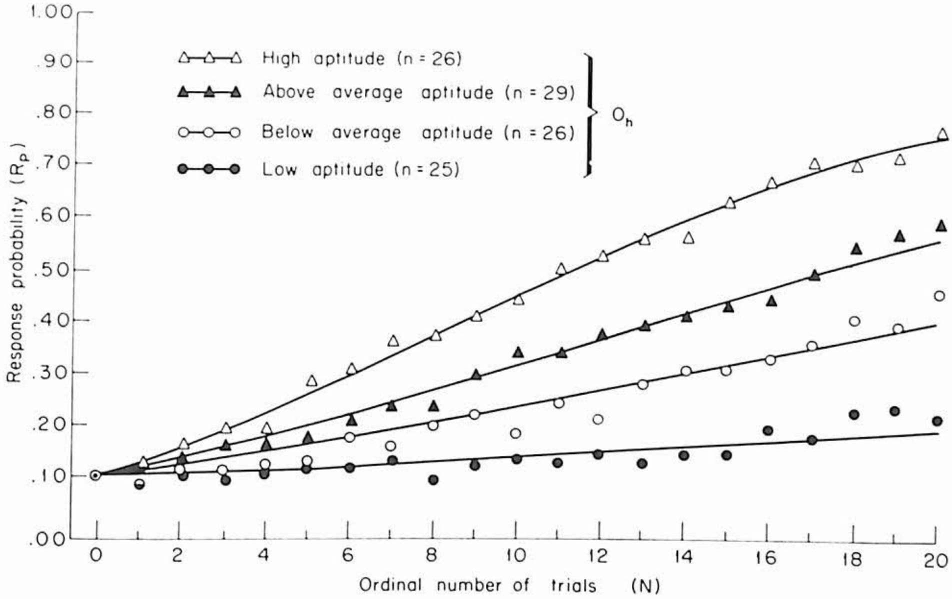


FIG. 8. Response probability curves ($R_p$) in selective learning as a function of the number of practice trials ($N$), with individual differences in aptitude ($O_h$) as the parameter. All curves have been fitted by Eq. (4), with only the $r$ parameter varying (after Noble *et al.*, 1958c, p. 164).

1965a) provided aptitude curves similar to those in Fig. 8. The role of the $O_h$ factor has also been explored with the Finger Maze. Commenting on his conventional serial-position effects for a nine-unit problem, Jones noted that the curves of four ability groups differed "quantitatively rather than qualitatively" (1945, p. 42). He, too, found no acquisition sex differences among 144 teenage boys and girls. For the Discrimination Reaction Timer, Adams (1957) published stable acquisition curves for 860 military trainees measured in terms of $R_t$ for 16 10-trial blocks of 3 seconds each. His first-rate discussion of the relationships among the concepts of abilities, aptitudes, capacities, factors, and learning operations should be required reading (along with the seminal papers of Woodrow and Hull cited earlier) for all activists who wish to do something constructive about the schism between experimental and differential psychology.

Moving on to a consideration of $R = f(O_s, O_y)$, but lingering awhile with the Discrimination Reaction Timer, I should not overlook the recent investigation of Noble, Baker, and Jones (1964). Quantitative relationships were sought among the variables $R_{1/t}$, $S_n$, $O_y$, and $O_s$ for 600 subjects in 30 groups between the ages of 8 and 87 years. All received 320 3-second trials (16 blocks of 20 trials each). Results should be exponential when the $R_{1/t} = f(N)$ curves are plotted for each group. Such was the case, felicitously, for Hull's theory. $H = M_H (1 - e^{-kN})$, plus an added constant, was found capable of describing all 30 acquisition curves with an average $F = 97.98\%$ when the asymptote, rate, and intercept parameters were varied jointly. Confirming and extending the classical age-performance data, acquisition speeds followed differential trends for both $O_y$ and $O_s$ variation while overall proficiency was a nonmonotonic function of $O_y$. There was a rapid growth to a maximum level for females at the age of 16 years, a later peak for males at 20 years, then came a gradual nonparallel decline into the seventh decade. Males performed significantly faster than females, and all two-factor interactions were significant. Age and sex, according to the Noble–Baker–Jones experiment, are critical parameters in selective learning and performance when $R_{1,t}$ scores are employed. Some important $O_s$ implications also arose in an experiment by Patel and Grant (1964). There was an interaction of $O_s$ with the $S_u$ and $S_n$ variables. Although there were no general sex differences during training, the rate of acquisition was slower for men initially and faster later as compared with women. As the authors put it, ". . . there was a clear-cut effort × distribution interaction for the women but not for the men" (1964, p. 229). Follow-up studies employing the Mathometer and other selective-learning apparatus may confidently be expected.

Results for the $O_d$ variable are mostly negative when (Taylor) anxiety scores are the defining criteria. Farber and Spence (1953) ran 80 Iowa students on a two-choice 10-unit stylus maze. In harmony with predictions from Eqs. (2) and (3), the high $O_d$ group (manifestly anxious by virtue of scoring above the fourth pentile) performed poorly relative to the low $O_d$ group (not manifestly anxious; below first pentile). Furthermore, at the most difficult choice points there occurred the greatest between-groups differences. Because the same high $O_d$ subjects were superior in eyelid conditioning, Farber and Spence understandably opted for the $D \times H$ hypothesis. A repetition upon 96 Rochester collegians by Axelrod, Cowen, and Heilizer (1956) failed to uncover any significant differences among three anxiety groups with respect either to total errors or trials-to-criterion scores. Grice (1955) sampled 300 military trainees, selecting 30 each from below the first and above the ninth anxiety deciles, then ran them on a Discrimination Reaction Timer modified to produce strong interference. Although his high $O_d$ group was significantly inferior in both $R_{1,t}$ and R− scores, Grice found a negative correlation between anxiety and intelligence. Covariance adjustments eradicated the performance differences. One of my doctoral students at Louisiana State tested 288 undergraduates on the Mathometer and failed to get any differences among three levels of $O_d$ (Brackenridge, 1956). There were no interactions of R+ or R− scores with $S_n$ and $S_c$ variables, although

Brackenridge's four-choice $R_p = f(N)$ curves were consistent with those reported earlier by Noble (1955) upon military trainees; the mean $F$ value for three groups whose $i$ (and $r$) parameters were .100 (.896), .167 (.881), and .250 (.842) was 99.60%. Similarly, in one of our experiments already cited there were notably low correlations between anxiety scores and scores on Finger Maze and Selective Mathometer. Our interpretation was that the tasks "were relatively non-stressful and thus failed to evoke different motivational levels in subjects whose anxiety scores varied" (Noble *et al.*, 1958c, p. 159).

Finally, there are significant main effects of social facilitation upon $R_{1/t}$ scores, as revealed by experiments with the Discrimination Reaction Timer (Noble, Fuchs, Robel, & Chambers, 1958d; Noble, Chambers, Fuchs, & Robel, 1959b; Noble, Fuchs, Walters, & Chambers, 1959c). For this reason it is important to control the number and relative positions of the subjects when multiple-unit apparatus are used.

## Summary

In this chapter I have attempted to provide a systematic analysis of the variables, theories, and laws of human nonverbal selective learning. This topical concentration has necessitated a de-emphasis of research upon perceptual skills, motor skills, information processing, classical verbal learning, Humphreys' verbal-expectation learning, problem solving, concept formation, predifferentiation transfer studies, engineering psychology, and experiments employing infrahuman subjects.

Multiple-choice situations were presented in historical perspective together with a survey and classification of six contemporary selective-learning apparatus.

The Hull–Spence theory was examined in detail and found to be experimentally indeterminate. To avoid this practical impasse I recommended the consideration of an approximation formula elaborated from the work of Gompertz and Courtis. The adequacy of the equation $R_p = a(i)^{r^N}$ was evaluated by curve-fitting techniques and its parameters were given preliminary rationalization. Predictability indices averaged 99.77% for three theoretical and three empirical sets of *group* data. A comparative analysis of 58,400 individual response predictions under three methods of practice on the Selective Mathometer utilizing the core hypotheses of postremity, frequency, and reinforcement indicated that the odds-on favorite is an S-R-contiguity–frequency-reinforcement theory which assumes an underlying continuity for the growth of associative strength. For intermediate stages of practice the average predictability of *individual* behavior is about 81%, but new developments in the Georgia laboratory offer the potential capability of achieving higher levels.

Finally, the laws of selective learning ($R_p = f[S_n]$) were tentatively codified under the rubrics of Task Complexity ($S_c$), Feedback Contingencies ($S_g$), Work Distribution ($S_w$), Transfer of Training ($S_t$), and Individual Differences ($O_h$, $O_s$, $O_y$, $O_d$). In each category the relevant literature was reviewed and a provisional coordination of operations with parameters was made in the approximation formula. Data generated from representative experimental variables are in close agreement with the equation, and its parameters vary according to theoretical expectation.

## References

Adams, J. A. The prediction of performance at advanced stages of training on a complex psychomotor task. *USAF Hum. Resour. Res. Cent. res. Bull.*, 1953, No. 53–49.

Adams. J. A. Effect of experimentally induced muscular tension on psychomotor performance. *J. exp. Psychol.*, 1954, 48, 127–130. (a)

Adams, J. A. Multiple versus single problem training in human problem solving. *J. exp. Psychol.*, 1954, **48**, 15–18. (b)

Adams, J. A. Psychomotor response acquisition and transfer as a function of control-indicator relationships. *J. exp. Psychol.*, 1954, **48**, 10–14. (c)

Adams, J. A. The relationship between certain measures of ability and the acquisition of a psychomotor criterion response. *J. gen. Psychol.*, 1957, **56**, 121–134.

Adams, J. A. Motor skills. *Annu. Rev. Psychol.*, 1964, **15**, 181–202.

Anderson, N. H. Effect of first-order conditional probability in a two-choice learning situation. *J. exp. Psychol.*, 1960, **59**, 73–93.

Anderson, N. H. An evaluation of stimulus sampling theory: comments on Professor Estes' paper. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 129–144.

Anderson, N. H., & Grant, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, **54**, 305–317.

Andreas, B. G. Empirical gradients of generalization in a perceptual-motor task. *J. exp. Psychol.*, 1954, **48**, 119–122.

Axelrod, H. S., Cowen, E. L., & Heilizer, F. The correlates of manifest anxiety in stylus maze learning. *J. exp. Psychol.*, 1956, **51**, 131–138.

Barch, A. M. Serial paired-associate performance, I: effect of sequence complexity and irrelevant cues. *Percept. mot. Skills*, 1960, **10**, 175–178. *(a)*

Barch, A. M. Toward an ideal perceptual-motor device. *Percept. mot. Skills*, 1960, **10**, 11–20. (b)

Besnard, G. G., Briggs, L. J., Mursch, G. A., & Walker, E. S. Development of the Subject-Matter Trainer. *USAF Personnel Train. Res. Center tech. Memo.*, 1955, No. 55–7.

Bilodeau, E. A. Acquisition of two lever-positioning responses practiced over several periods of alternation. *J. exp. Psychol.*, 1953, **46**, 43–49.

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Bilodeau, E. A., Brown, J. S., & Meryman, J. J. The summation of generalized reactive tendencies. *J. exp. Psychol.*, 1956, **51**, 293–298.

Birch, D. Incentive value of success and instrumental approach behavior. *J. exp. Psychol.*, 1964, **68**, 131–139.

Birch, D., & Vandenberg, V. The necessary conditions for cue-position patterning. *J. exp. Psychol.*, 1955, **50**, 391–396.

Bower, G. H. Choice-point behavior. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press., 1959. Pp. 109–124.

Brackenridge, C. W. *The role of manifest anxiety in complex learning.* Unpublished doctoral dissertation. Louisiana State Univer., 1956.

Brogden, W. J. Acquisition and extinction of a conditioned avoidance response in dogs. *J. comp. physiol. Psychol.*, 1949, **42**, 296–302.

Brogden, W. J., & Schmidt, R. E. Acquisition of a 24-unit verbal maze as a function of number of alternate choices per unit. *J. exp. Psychol.*, 1954, **48**, 335–338. (a)

Brogden, W. J., & Schmidt, R. E. Effect of number of choices per unit of a verbal maze on learning and serial position errors. *J. exp. Psychol.*, 1954, **47**, 235–240. (b)

Brown, J. S., Bilodeau, E. A., & Baron, M. R. Bidirectional gradients in the strength of a generalized voluntary response to stimuli on a visual-spatial dimension. *J. exp. Psychol.*, 1951, **41**, 52–61.

Brown, J. S., Clarke, F. R., & Stein, L. A new technique for studying spatial generalization with voluntary responses. *J. exp. Psychol.*, 1958, **55**, 359–362.

Bush, R. R., & Estes, W. K. (Eds.) *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press, 1959.

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Cantor, G. N., & Spiker, C. C. Effects of nonreinforced trials on discrimination learning in preschool children. *J. exp. Psychol.*, 1954, **47**, 256–258.

Cantor, J. H. Amount of pretraining as a factor in stimulus predifferentiation and performance set. *J. exp. Psychol.*, 1955, **50**, 180–184.

Chambers, R. W., & Noble, C. E. Effects of various asymptotic restrictions on human trial-and-error learning. *J. exp. Psychol.*, 1961, **61**, 417–421.

Courtis, S. A. *The measurement of growth.* Ann Arbor, Mich.: Brumfield & Brumfield, 1932.

Deese, J. *The psychology of learning* (2nd ed.). New York: McGraw-Hill, 1958.

Denny, M. R., & Dunham, M. D. The effect of differential non-reinforcement of the incorrect response on the learning of the correct response in the simple T-maze. *J. exp. Psychol.*, 1951, **41**, 382–389.

Duncan, C. P. Transfer in motor learning as a function of degree of first-task learning and inter-task similarity. *J. exp. Psychol.*, 1953, **45**, 1–11.

Duncan, C. P. Development of response generalization gradients. *J. exp. Psychol.*, 1955, **50**, 26–30.

Duncan, C. P. Transfer after training with single versus multiple tasks. *J. exp. Psychol.*, 1958, **55**, 63–72.

Duncan, C. P., & Underwood, B. J. Retention of transfer in motor learning after twenty four hours and after fourteen months. *J. exp. Psychol.*, 1953, **46**, 445–452.

Edwards, W. Probability learning in 1000 trials. *J. exp. Psychol.*, 1961, **62**, 385–394.

Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: a study of a science.* Vol. 2. New York: McGraw-Hill, 1959. Pp. 380–491.

Estes, W. K. Learning theory. *Annu. Rev. Psychol.*, 1962, **13**, 107–144.

Estes, W. K. Probability learning. In A. W. Melton (Ed.), *Categories of human learning.* New York: Academic Press, 1964. Pp. 89–128.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, **47**, 225–234.

Farber, I. E., & Spence, K. W. Complex learning and conditioning as a function of anxiety. *J. exp. Psychol.*, 1953, **45**, 120–125.

Farese, F. J., & Noble, C. E. Trial-and-error vs. mixed-selective learning in man. *Percept. mot. Skills*, 1960, **10**, 115–122.

Fitts, P. M. Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning.* New York: Academic Press, 1964. Pp. 243–285.

Fleishman, E. A., & Hempel, W. E., Jr. Changes in factor structure of a complex psychomotor task as a function of practice. *Psychometrika*, 1954, **19**, 239–252.

Fleishman, E. A., & Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. *J. exp. Psychol.*, 1955, **49**, 301–312.

Fuchs, J. E., Leland, Q. A., & Noble, C. E. Compound trial-and-error learning under delayed terminal reward and response competition. *Proc. Mont. Acad. Sci.*, 1960, **20**, 91–92.

Gerjuoy, I. R., Gerjuoy, H., & Mathias, R. Probability learning: left-right variables and response latency. *J. exp. Psychol.*, 1964, **68**, 344–350.

Gompertz, B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of live contingencies. *Phil. Trans.* 1825, **115**, Pt. 1, 513–585.

Grant, D. A., Hornseth, J. P., & Hake, H. W. The influence of the inter-trial interval on the Humphreys' 'random reinforcement' effect during the extinction of a verbal response. *J. exp. Psychol.*, 1950, **40**, 609–612.

Grice, G. R. Discrimination reaction time as a function of anxiety and intelligence. *J. abnorm. soc. Psychol.*, 1955, **50**, 71–74.

Guthrie, E. R. *The psychology of learning* (rev. ed.) New York: Harper, 1952.

Guthrie, E. R. Association by contiguity. In S. Koch (Ed.), *Psychology: a study of a science.* Vol. 2. New York: McGraw-Hill, 1959. Pp. 158–195.

Hake, H. W., Grant, D. A., & Hornseth, J. P. Resistance to extinction and the pattern of reinforcement, III: the effect of trial patterning in verbal 'conditioning.' *J. exp. Psychol.*, 1951, **41**, 221–225.

Hamilton, G. B. A study of trial and error reactions in mammals. *J. anim. Behav.*, 1911, **1**, 33–66.

Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: a study of a science.* Vol. 2. New York: McGraw-Hill, 1959. Pp. 492–537.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 517–567.

Hilgard, E. R. *Theories of learning* (2nd ed.). New York: Appleton, 1956.

Hubbard, W. R. Secondary reinforcement of a simple discrimination in human beings. *J. exp. Psychol.*, 1951, **41**, 233–241.

Hull, C. L. Simple trial-and-error learning: a study in psychological theory. *Psychol. Rev.*, 1930, **37**, 241–256.

Hull, C. L. *Principles of behavior.* New York: Appleton, 1943.

Hull, C. L. The place of innate individual and species differences in a natural-science theory of behavior. *Psychol. Rev.*, 1945, **52**, 55–60.

Hull, C. L. *A behavior system.* New Haven, Conn.: Yale Univer. Press, 1952.

Hull, C. L., & Spence, K. W. 'Correction' vs. 'non-correction' method of trial-and-error learning in rats. *J. comp. Psychol.*, 1938, **25**, 127–145.

Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, **25**, 294–301.

Irion, A. L., & Briggs, L. J. Learning task and mode of operation variables in use of Subject-Matter Trainer. *USAF Personnel Train. Res. Center tech. Rep.*, 1957, No. 57-8.

Jeffrey, W. E., & Cohen, L. B. Effect of spatial separation of stimulus, response, and reinforcement on selective learning in children. *J. exp. Psychol.*, 1964, **67**, 577–580.

Jones, H. E. Trial and error learning with differential cues. *J. exp. Psychol.*, 1945, **35**, 31–45.

Kaess, W., & Zeaman, D. Positive and negative knowledge of results on a Pressey-type punchboard. *J. exp. Psychol.*, 1960, **60**, 12–17.

Kendler, H. H. Learning. *Annu. Rev. Psychol.*, 1959, **10**, 43–88.

Kimble, G. A. *Hilgard and Marquis' conditioning and learning* (rev. ed.). New York: Appleton, 1961.

Lewis, D. *Quantitative methods in psychology.* New York: McGraw-Hill, 1960.

Lewis, D., & Miles, G. H. Retroactive interference in performance on the Star Discrimeter as a function of amount of interpolated learning. *Percept. mot. Skills*, 1956, **6**, 295–298.

Lewis, D., McAllister, D. E., & Adams, J. A. Facilitation and interference in performance on the modified Mashburn apparatus, I: the effects of varying the amount of original learning. *J. exp. Psychol.*, 1951, **41**, 247–260.

Loess, H. B., & Duncan, C. P. Human discrimination learning with simultaneous and successive presentation of stimuli. *J. exp. Psychol.*, 1952, **44**, 215–221.

McAllister, D. E. The effects of various kinds of relevant verbal pretraining on subsequent motor performance. *J. exp. Psychol.*, 1953, **46**, 329–336.

McAllister, D. E., & Lewis, D. Facilitation and interference in performance on the modified Mashburn apparatus, II: the effects of varying the amount of interpolated learning. *J. exp. Psychol.*, 1951, **41**, 356–363.

McFann, H. H. Effects of response alteration and different instructions on proactive and retroactive facilitation and interference. *J. exp. Psychol.*, 1953, **46**, 405–410.

McGeoch, J. A., & Irion, A. L. *The psychology of human learning* (2nd ed.). New York: Longmans, 1952.

Mandler, G. Transfer of training as a function of degree of response overlearning. *J. exp. Psychol.*, 1954, **47**, 411–417.

Mandler, G., & Kuhlman, C. K. Proactive and retroactive effects of overlearning. *J. exp. Psychol.*, 1961, **61**, 76–81.

Marx, M. H. Gradients of error-reinforcement in a serial perceptual-motor task. *Psychol. Monogr.*, 1957, **71**, No. 8 (Whole No. 437). (a)

Marx, M. H. Gradients of error reinforcement in normal multiple-choice learning situations. *J. exp. Psychol.*, 1957, **54**, 225–228. (b)

Melton, A. W. (Ed.) *Apparatus tests.* Washington, D.C.: U.S. Govt. Printing Office, 1947. (AAF Aviat. Psychol. Prog. Res. Rep. No. 4).

Montgomery, V. E. Transfer of training in motor learning as a function of distribution of practice. *J. exp. Psychol.*, 1953, **46**, 440–444.

Morin, R. E., & Grant, D. A. Learning and performance on a key-pressing task as function of the degree of spatial stimulus-response correspondence. *J. exp. Psychol.*, 1955, **49**, 39–47.

Morrisett, L., Jr., & Hovland, C. I. A comparison of three varieties of training in human problem solving. *J. exp. Psychol.*, 1959, **58**, 52–55.

Murphy, J. V., & Miller, R. E. Spatial contiguity of cue, reward, and response in discrimination learning by children. *J. exp. Psychol.*, 1959, **58**, 485–489.

Neimark, E. D. Effects of type of non-reinforcement and number of available responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, **52**, 209–220.

Noble, C. E. An estimation of the absolute difficulties of component tasks in the Complex Coordination Test (CM 701 E). *USAF Hum. Resources Res. Cent. Res. Note*, 1952, No. 52–8. (a)

Noble, C. E. A preliminary analysis of the relative difficulties of component tasks in the Complex Coordination Test (CM 701 E). *USAF Hum. Resources Res. Cent. Res. Note*, 1952, No. 52–6. (b)

Noble, C. E. A representative scale of difficulty in the Complex Coordination Test (CM 701 E). *USAF Hum. Resources Res. Cent. Res. Bull.*, 1952, No. 52–41. (c)

Noble, C. E. Amount set and the length-difficulty function for a self-paced perceptual-motor skill. *J. exp. Psychol.*, 1953, **46**, 435–439. (a)

Noble, C. E. Some physical sources of difficulty in the Complex Coordination Test (CM 701 E). *USAF Hum. Resources Res. Cent. Res. Bull.*, 1953, No. 53–7. (b)

Noble, C. E. The difficulty-complexity relationship in compound trial-and-error learning. *Amer. Psychologist*, 1954, **9**, 443.

Noble, C. E. Compound trial-and-error learning as a function of response availability ($N_R$). *J. exp. Psychol.*, 1955, **49**, 93–96.

Noble, C. E. Human trial-and-error learning. *Psychol. Rep.*, 1957, **3**, 377–398 (Monogr. Suppl. 8). (a)

Noble, C. E. Supplementary report: response availability in human trial-and-error learning. *J. exp. Psychol.*, 1957, **53**, 359–360. (b)

Noble, C. E. The length-difficulty relationship in compound trial-and-error learning. *J. exp. Psychol.*, 1957, **54**, 246–252. (c)

Noble, C. E. Predicting human choice-point behavior in selective mathometry. Paper given at convention of Sth. Soc. for Phil. and Psychol., Memphis, April 1962.

Noble, C. E. Human selective learning under correction and non-correction methods. Paper given at convention of Midw. Psychol. Ass., Chicago, May 1965. (a)

Noble, C. E. New studies on the prediction of individual behavior in human trial-and-error learning. Paper given at convention of Psychonomic Soc., Chicago, October 1965. (b)

Noble, C. E. Partial vs. continuous reinforcement in human selective learning. Paper given at convention of Sth. Soc. for Phil. and Psychol., Atlanta, April 1965. (c)

Noble, C. E. Types and combinations of feedback in human trial-and-error learning. Paper given at convention of Sthea. Psychol. Ass., Atlanta, April 1965. (d)

Noble, C. E., & Alcock, W. T. Human delayed-reward learning with different lengths of task. *J. exp. Psychol.*, 1958, **56**, 407–412.

Noble, C. E., & Farese, F. J. An apparatus for research in human selective learning. *J. Psychol.*, 1955, **39**, 475–484.

Noble, C. E., & Fuchs, J. E. Response availability and task length in human complex learning. Paper given at convention of Psychonomic Soc., New York, August, 1961.

Noble, C. E., & Noble, J. L. Human trial-and-error learning under joint variation of locus of reward and type of pacing. *J. exp. Psychol.*, 1958, **56**, 103–109.

Noble, C. E., & Taylor, A. Influence of work distribution upon complex learning by the noncorrection and modified-correction methods. *J. exp. Psychol.*, 1959, **58**, 352–356.

Noble, C. E., Alcock, W. T., & Frye, R. L., Jr. The joint influence of practice and instructions on discrimination reaction time. *J. Psychol.*, 1959, **48**, 125–130. (a)

Noble, C. E., Chambers, R. W., Fuchs, J. E., & Robel, D. P. A further study of individual vs. social performance in discrimination reaction. *Proc. Mont. Acad. Sci.*, 1959, **19**, 150. (b)

Noble, C. E., Fuchs, J. E., Walters, H. A., & Chambers, R. W. One-man vs. two-man performance on the discrimination reaction time test. *Proc. Mont. Acad. Sci.*, 1959, **19**, 151. (c)

Noble, C. E., Fuchs, J. E., & Thompson, J. D. A new automatic circuit for the Selective Mathometer (MK II). *J. Psychol.*, 1963, **55**, 241–257.

Noble, C. E., Alcock, W. T., & Farese, F. J. Habit reversal under differential instructions in compound trial-and-error learning. *J. Psychol.*, 1958, **46**, 253–264. (a)

Noble, C. E., Alcock, W. T., & Noble, J. L. The interaction of type of choice procedure with amount of practice in trial-and-error learning under two reward conditions. *J. Psychol.*, 1958, **46**, 295–301. (b)

Noble, C. E., Noble, J. L., & Alcock, W. T. Prediction of individual differences in human trial-and-error learning. *Percept. mot. Skills*, 1958, **8**, 151–172 (Monogr. Suppl. 2). (c)

Noble, C. E., Fuchs, J. E., Robel, D. P., & Chambers, R. W. Individual vs. social performance on two perceptual-motor tasks. *Percept. mot. Skills*, 1958, **8**, 131–134. (d)

Noble, C. E., Baker, B. L., & Jones, T. A. Age and sex parameters in psychomotor learning. *Percept. mot. Skills*, 1964, **19**, 935–945.

Nystrom, C. O., & Grant, D. A. Performance on a key pressing task as a function of the angular correspondence between stimulus and response elements. *Percept. mot. Skills*, 1955, **5**, 113–125 (Monogr. Suppl. 1).

Overall, J. E., & Brown, W. L. Recency, frequency, and probability in response prediction. *Psychol. Rev.*, 1957, **64**, 314–322.

Overall, J. E., & Brown, W. L. Cognitive-postremity predictions of learning behavior. *J. gen. Psychol.*, 1960, **62**, 189–197.

Patel, A. S., & Grant, D. A. Decrement and recovery effects in a perceptual-motor learning task as a function of effort, distribution of practice, and sex of subject. *J. gen. Psychol.*, 1964, **71**, 217–231.

Paul, N. T., & Noble, C. E. Influence of successive habit reversals on human learning and transfer. *J. exp. Psychol.*, 1964, **68**, 37–43.

Perkins, M. J., Banks, H. P., & Calvin, A. D. The effect of delay on simultaneous and successive discrimination in children. *J. exp. Psychol.*, 1954, **48**, 416–417.

Peterson, J. Learning when frequency and recency factors are negative. *J. exp. Psychol.*, 1922, **5**, 270–300.

Prentice, W. C. H. Continuity in human learning. *J. exp. Psychol.*, 1949, **39**, 187–194.

Restle, F. A survey and classification of learning models. In R. R. Bush & W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford, Calif.: Stanford Univer. Press, 1959. Pp. 415–427.

Restle, F. *Psychology of judgment and choice: a theoretical essay*. New York: Wiley, 1961.

Restle, F. The selection of strategies in cue learning. *Psychol. Rev.*, 1962, **69**, 329–343.

Riley, D. A. Rote learning as a function of distribution of practice and the complexity of the situation. *J. exp. Psychol.*, 1952, **43**, 88–95.

Seibel, R. Discrimination reaction time as a function of the number of stimulus-response pairs and the self-pacing adjustment of the subject. *Psychol. Monogr.*, 1962, **76**, No. 42 (Whole No. 561).

Seibel, R. Discrimination reaction time for a 1,023-alternative task. *J. exp. Psychol.*, 1963, **66**, 215–226.

Seward, J. P., Dill, J. B., & Holland, M. A. Guthrie's theory of learning: a second experiment. *J. exp. Psychol.*, 1944, 34, 227–238.

Shock, N. W., & Morales, M. F. A fundamental form for the differential equation of colonial and organism growth. *Bull. math. Biophys.*, 1942, 4, 63–71.

Simon, H. A. Amounts of fixation and discovery in maze learning behavior. *Psychometrika*, 1957, **22**, 261–268.

Skinner, B. F. *Cumulative record* (enlarged ed.). New York: Appleton, 1961.

Spence, K. W. The reliability of the maze and methods of its determination. *Comp. Psychol. Monogr.*, 1932, 8, 1–45.

Spence, K. W. *Behavior theory and conditioning.* New Haven, Conn.: Yale Univer. Press, 1956.

Spence, K. W. *Behavior theory and learning.* Englewood Cliffs, N. J.: Prentice-Hall, 1960.

Spiker, C. C., & Holton, R. B. Associative transfer in motor paired-associate learning as a function of amount of first-task practice. *J. exp. Psychol.*, 1958, **56**, 123–132.

Taylor, A., & Noble, C. E. Acquisition and extinction phenomena in human trial-and-error learning under different schedules of reinforcing feedback. *Percept. mot. Skills*, 1962, **15**, 31–44.

Terrell, G., Jr., & Kennedy, W. A. Discrimination learning and transposition in children as a function of the nature of the reward. *J. exp. Psychol.*, 1957, **53**, 257–260.

Thorndike, E. L. Animal intelligence. An experimental study of the associative processes in animals. *Psychol. Monogr.*, 1898, **2**, No. 8.

Thurstone, L. L. The relation between learning time and length of task. *Psychol. Rev.*, 1930, **37**, 44–53.

Tolman, E. C. *Purposive behavior in animals and men.* New York: Century, 1932.

Underwood, B. J. Learning. *Annu. Rev. Psychol.*, 1953, 4, 31–58.

Underwood, B. J., & Keppel, G. One-trial learning? *J. verb. Learn. verb. Behav.*, 1962, **1**, 1–13.

Voeks, V. W. Postremity, recency, and frequency as bases for prediction in the maze situation. *J. exp. Psychol.*, 1948, **38**, 495–510.

Walk, R. D. Effect of discrimination reversal on human discrimination learning. *J. exp. Psychol.*, 1952, **44**, 410–419.

Waters, R. H., & Reitz, J. G. The role of recency in learning. *J. exp. Psychol.*, 1950, **40**, 254–259.

White, B. N., & Spiker, C. C. The effect of stimulus similarity on amount of cue-position patterning in discrimination problems. *J. exp. Psychol.*, 1960, **59**, 131–136.

Witte, R. S. Long-term effects of patterned reward schedules. *J. exp. Psychol.*, 1964, **68**, 588–594.

Woodrow, H. The effect of practice on groups of different initial ability. *J. educ. Psychol.*, 1938, **29**, 268–278.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology* (rev. ed.). New York: Holt, 1954.

Wyckoff, L. B., & Sidowski, J. B. Probability discrimination in a motor task. *J. exp. Psychol.*, 1955, **50**, 225–231.

Yerkes, R. M. The study of human behavior. *Science*, 1914, **39**, 625–633.

Yerkes, R. M. The mental life of monkeys and apes: A study of ideational behavior. *Behav. Monogr.* 1916, **3**, 1–145.

Yerkes, R. M. A new method of studying the ideational behavior of mentally defective and deranged as compared with normal individuals. *J. comp. Psychol.*, 1921, **1**, 369–394.

Zeiler, M. D. Component and configurational learning in children. *J. exp. Psychol.*, 1964, **68**, 292–296. (a)

Zeiler, M. D. Transposition in adults with simultaneous and successive stimulus presentation. *J. exp. Psychol.*, 1964, **68**, 103–107. (b)

# Selective Learning
## Comments on Professor Noble's Paper

DAVID BIRCH
*University of Michigan*

Tolman in his presidential address to the American Psychological Association in 1937 said, "Let me close, now, with a final confession of faith. I believe that everything important in psychology (except perhaps such matters as the building up of a superego, that is everything save such matters as involve society and words) can be investigated in essence through the continued experimental and theoretical analysis of the determiners of rat behavior at a choice point in a maze." (Tolman, 1938, p. 34). Though I think this is probably more true than false, I don't think that it is wholly true. Nor do I think all of the important problems of selective learning in human beings can be solved using the Multiple Serial Discrimeter and the Selective Mathometer, at least not in their present form and in terms of their present usage. It is certainly true, however, that, in the hands of Dr. Noble and others he refers to in his chapter, these devices have provided a large quantity of systematic data, and we all agree that data are the foundation that must support any theory building. This last, commonly accepted thought carries certain undesirable possibilities for all of us, however, for just as we tend to write our theories in terms of the data we have so we also tend to build our apparatus and design our experiments in terms of our theories. The combination of these two events can lead us to data and conceptualizations that are of lesser scope than we intend or wish. My discussion is derived from the concern implied in these comments and is intended to foster disquiet about our current methods and theories in the complex domain of human selective learning.

## Theory of Selective Learning

Noble approaches the problem of selective learning from the standpoint of Hull–Spence theory. This theory calls on a number of intervening variables, only one of which is $H$. Each variable does its part in the division of labor and does so in a highly organized and systematic way. Unfortunately, the statements of this organization among variables that link observable S to observable R become very unwieldy mathematically. Thus, the Hull–Spence theory, while preserving all the variables in their proper relationships with each other, does not yield an algebraic solution for probability of correct response $(R_p)$ as a function of the number of trials $(N)$.

Noble's solution to this unsatisfactory state of affairs is to offer an approximation formula, $R_p = a(i)^{r^N}$. This expression was chosen to meet six requirements set for it (see p. 60), and it seems to do so admirably. Somehow in this procedure, however, the variables of the Hull–Spence theory got lost or collapsed or reorganized so as to be missing in the approximation formula. Somewhere along the line Noble's equation severed its connections to the roots from which it sprung.

It seems to me that Noble has not offered an approximation formula for the Hull–Spence theory but is making a proposal that is just plain different from it except in general emphasis and orientation. If that be granted, let us ask what sort of theory it is and what kind of job it is doing. The double-exponential growth curve along with the experimental identifications of the parameters summarizes the theory. It is a statement intended to represent the data from selective learning experiments and to provide a framework for the inclusion of the effects of additional variables which might be incorporated into selective learning. It would appear to accomplish both aims with distinction. For example, Noble's own research program has included manipulations of response availability, temporal locus of reward, type of pacing, habit reversal, work distribution, and certain individual difference parameters, and the basic equation was used to interpret their effects. Additionally, the utility of the equation in summarizing the gross features of the data, as indicated by the degree of curve fitting obtained, is quite clear. The theory does not appear to be well suited, however, to come to grips with additional aspects of the data such as variances, trials to criterion, trials to the $K$th success, run characteristics, and other statistics considered important by other mathematically oriented theorists (e.g., Bush & Estes, 1959).

Leaving aside for the moment the question of what kind and how much data a theory should wrestle with, I think we might ask where we should go next with selective learning. What additional directions are there to be taken and what new questions might be asked? Fundamentally, we are interested in two things; i.e., there are two things we would like information about and theory for. These are (1) the dynamic process occurring on a single instance in which a selection is made and (2) the fate of the variables determining the dynamic process as repeated instances of selection occur.

The dynamic process of selection has received rudimentary treatment by Estes (1950), who adapted probability theory to his conception of stimulus sampling, and by Spence (1954), who utilized the same probability theory in conjunction with the Hullian vocabulary. In both, the dynamic process conceived to operate on the occasion of selection is randomly based. For Estes, stimulus sampling is repeated until a stimulus element connected to the appropriate response is obtained and for Spence, oscillation of effective excitatory potential continues until a superthreshold value is garnered. Such models of the selection process lead to predictions concerning both the choice of alternative and the latency of the choice. More recent conceptions (e.g., La Berge, 1962) have increased the complexity, if not the determinism, of the hypothesized processes.

The second point of interest, what happens to the relevant variables as repeated selections occur, is what is commonly captured in theories of learning. What I have called repeated instances of selection is what is usually called trials, and certain effects of these trials are ordinarily termed learning. I would like to suggest that there may be dividends to be gained by attempting to account for selective learning in terms of changes in the variables entering into the dynamic process of selection within a trial instead of our customary approach of attempting to account for the selection within a trial in terms of hypothesized learning processes cumulating over trials. In the present discussion, however, I wish to focus on the first point of interest, the dynamic process within a trial, as in greater need of attention.

## An Expanded Domain for Selective Learning

### ONGOING ACTIVITY

The study of human selective learning is a part of the study of action or, more precisely, it is the study of certain changes in a person's activity. In the phrase "changes in a person's activity" there is implicit recognition that the individual is engaging in some sort of ongoing activity at the time that a signal for a new activity appears. I want to make this recognition explicit and begin to explore some consequences this might have for selective learning experimentation and theory. Change in activity is possible only if activity is in progress and our theories should reflect this fact. In their general setting they should capture the idea of behavior as a continuously shifting set of activities not as a series of discrete, discontinuous events. This position is given an eloquent treatment by John W. Atkinson in his recent book *An Introduction to Motivation* (1964) under the topic "A New Premise for the Study of Motivation."

What, we may ask, however, is the ongoing activity in selective learning? In order to answer this I would like to make an imaginary modification in the Selective Mathometer. Think of the subject placing his hand on a knob which is the control device for some tracking task. He might, for example, be able to turn the knob like a steering wheel and keep a marker on a target. The usual Selective Mathometer task requires a response incompatible with such knob turning and is imposed on this continuing background activity. Immediately we wonder what will happen when a stimulus for selective learning is presented. Will the subject replace his knob turning with a key press or might he continue with the tracking task? Certainly it is not guaranteed that he will key press, for it should be possible to make the ongoing activity so attractive, perhaps by paying money, that the subject wouldn't undertake the selective learning at all.

But let's assume the subject does perform a key press each time a stimulus comes on and ask whether his performance in selective learning might differ as a function of the character and attractiveness of the ongoing tracking activity. It seems reasonable to think it might.

The fact that the subject is tracking at the time the new stimulus comes on must mean that tracking is supported by the strongest tendency in the subject's hierarchy. By the same thinking it must mean that another tendency, one to press a particular key, has gained dominance when behavior shifts. Further, the latency of the key pressing response is the period of time that passes before the new tendency dominates the old. Clearly, there must have been a shift in the hierarchy of tendencies *within* that time period. Such a shift could occur if (1) the tracking tendency were constant but the key pressing tendency became stronger or (2) the tracking tendency diminished in strength and the key pressing tendency became stronger, remained constant, or even diminished at a lesser rate or (3) the tracking tendency gained strength but the key pressing tendency gained even greater strength. Any of these dynamic processes would result in the behavior shift indicated.

Under the rule that the behavior of an organism at any given point in time is determined by the strongest tendency present, the selection of one key over another is governed by the speed with which the tendencies change their strengths. If key 4 is selected, it means the tendency associated with that key gained dominance over the tendency supporting the ongoing tracking. The tendencies for the other keys did not. The subject stopped tracking and pressed key 4 as soon as the latter tendency dominated the former. The tendency to press key 4 also dominated the tendencies for the other keys, but this is somewhat incidental since the shift in activity is defined with reference to the ongoing tracking. With this kind of analysis the theoretical focus is on the dynamics of tendency change within the time period of a trial and experimental questions center at the factors that are effective in changing strengths of tendencies.

What might happen to performance on the Selective Mathometer if an attractive ongoing activity were introduced? First, latencies of key pressing should be higher since the strength of the tendency for the ongoing behavior is greater and more time should be required to dominate it. Second, it would appear, at least on superficial examination, that the proportion of choices of keys would be unaffected unless there were special dependencies between the key responses and the tracking. On the other hand, it might take more trials to reach criterion because the experimental factors have their influence attenuated, for example, by preventing between-trials rehearsal.

This account is admittedly sketchy and speculative, but viewed this way, past research on selective learning has been confined largely to the very special case where the rôle of the ongoing activity is kept to a minimum. If the setting for selective learning is expanded to include explicit recognition of ongoing activity then an expanded theory is needed. Such a theory will deal with the basic question of what conditions need obtain in order for a shift in behavior to occur at all, as well as the question of which alternative will be selected. I wish I were able to report data pertinent to this kind of theory to you, but I can't. So far our success in this area has been greater with rats than with human subjects, and this is not the place to report rat data. In summary, then, it seems to me that if we take our definitions of response seriously we are led quite naturally and in a straightforward manner to a broadened conception of selective learning. This conception emphasizes the dynamic processes occurring within a trial and relates performance in selective learning to characteristics of ongoing as well as alternative activities.

### VERBAL CONTROL

Concern for the response in human selective learning leads in another direction. Put forthrightly, this direction is toward explicating the behavior process, including contributions from the verbal mechanism if appropriate, that takes place between receipt of a stimulus and the production of a response. Any theory in psychology aims at making valid statements about responses, and a theorist is free to choose the responses that interest him or that he thinks are important. Various effects on responses are sought out empirically, and one level of theorizing takes on the task of summarizing and coordinating these effects. Theory at this level does not attempt to answer the question of how or by what process a given variable exerts its effect. I think Noble's theory, as captured by the expression $R_p = a(i)^{r \cdot N}$, is an example here.

There is another level of theory that does include statements of process. This kind of theory may end up with a mathematical statement relating R to S, but it does so as a result of following through the processes of the theory. In this case the emphasis is on the effective agents operating within the theory that are responsible for the effects observed. This is the context into which I wish to insert a discussion of verbal control.

The main premises of my discussion are that the verbal mechanism is often an important contributor to the behavior the experimenter observes in his subject and that our theories in general do not take this into account. This is the kind of point made elsewhere by Dulany (1961), Farber (1963), and Birch (1964).

Consider the subject in a selective learning experiment. He is seated, instructed, and turned loose on the task. Now, he may be uncooperative, perhaps carelessly or perhaps maliciously, and all the experimenter can do is hope he is able to tell it so that he can throw the data out. But there is a deeper and theoretically more interesting issue involved than just identifying errant subjects. The fact that experimenters can encounter uncooperative subjects should in itself raise the sobering questions of which determinants of the behavior are attributable to intended manipulations in the experiment and which are not and to what extent do our theories represent adequately the intervening processes of behavior, particularly as they might involve the verbal mechanism.

To look further into this let us assume a cooperative subject, but one who might talk to himself. Behavior theories don't often make a place for a subject talking to himself, and I suspect the reason is that many of our theories for human beings are derived from theories for rats, and the issue of rat's talking to themselves seldom comes up. The point of concern here is not whether verbal responses obey different laws from other responses (as far as I know we have no reason to think they do) but rather whether a theory of the processes bridging from stimulus to response in an articulate organism, known to provide self-instructions on occasion, can be satisfactory without including a place for verbal control. I submit that a subject's talking to himself is a crucial part of his behavior in many situations. It may not be the part the experimenter is observing, it may be determined by variables the experimenter has not controlled, and it may give rise to additional determinants of the observed behavior that the experimenter is not counting on but, in general, it must be occurring.

There are several things a subject might talk to himself about in selective learning. He might rehearse between trials. This is not so bad because it can just be included within the definition of a trial although certainly part of the total response process is unspecified, and one trial may not be very much like another if they can differ in rehearsal. The subject might discuss with himself something about the order in which the stimuli occur. He might wonder if the experimenter is changing the answers on him. Finally, and most importantly for my present concerns, the subject might plan ahead. That is, he may not behave in terms of discrete S-R trials as carefully designed by the experimenter, but he may decide on trial $n$ what he will do on trial $n + j$ and perhaps on all intermediate trials. The reason this matters is that the actual motor response as observed and recorded is only a running off of the consequences of the verbalizations. It was the verbal response that was affected by the variables of the experiment, yet it is the motor response that is under treatment by the theory.

This kind of planning ahead on the part of a subject must force the theorist either to reconsider the definition of the response he is trying to account for or to re-examine his catalog of determinants of the response he is observing. The problem is even more serious if the theorist is primarily concerned with the process itself and is attempting a construction to that end.

The most extensive and systematic theoretical work on the role of verbal control in human learning situations has been that of Dulany (1962). He has faced the problem head-on by writing a theory around the contributions of the verbal mechanism. Historically, there has been another method of approach; namely, to devise experimental situations in which verbal control of the response is thought not to occur, at least not directly. Studied under the label of learning without awareness, this domain has been a morass that a number have entered only to return wearied and thankful to get back onto apparently firmer theoretical and methodological ground. Nevertheless, this topic seems to me to cut to the heart of theorizing about human behavior and at the risk of injecting results into a discussion section, I would like to mention two of our experimental attempts.

With our method the subject performs a task that is meaningful to him but which also allows us to measure aspects of his behavior that he is apparently not aware of. Specifically, we measure the speed with which the subject reaches for his materials. In one experiment cited by Noble these were puzzles which the subject was to solve and in another, more recent experiment, beads which signified gains or losses of money. For the latter experiment, in which I was assisted by Elizabeth Ligon, the arrangement for timing was improved by employing a proximity switch. On each trial the subject pushed a button to indicate whether he thought he would win or lose and then reached into a container of 100 beads, 50 of which were white and 50 colored. Three containers, yellow, red and blue, matching the colored beads they held and signifying losses of 1, 5, or 10 cents, were presented equally often. A white bead uniformly won 5 cents. A clock started when the subject pushed the win or lose button and stopped when he reached into the container. The subject was not told he was being timed, and he had no reason to suspect it since he thought the button indicated his win-lose judgment and since the proximity switch stopped the clock without his knowledge. After 173 trials the experimenter changed one of the containers from 50% to 0% win without the subject's knowing and gave 81 additional trials. No subject showed awareness of being timed when interviewed subsequently.

For 38 male undergraduate subjects two main results were obtained: (1) an inverse relation of speed of reaching to magnitude of loss during preshift trials and (2) a decline in the speed of reaching to the container shifted to 0% win. These data show that speed of reaching, apparently not under the direct verbal control of the subject, is influenced by the incentive value of the outcome. The delicate question of whether the reaching is under indirect verbal control; i.e., is governed by direct verbal control over some correlated behavior, is not answered by these experiments. Certainly, however, the control is different from that which would have been exerted had we asked the subject to indicate the degree of his anticipation of a negative outcome by the speed with which he moved his hand, as we might have asked in a psychophysical scaling experiment.

It is more difficult to get satisfactory removal of direct verbal influence in selective learning situations of the classical form. In daily life there appear to be numerous occasions in which we make selections without an awareness of using intervening verbal processes, but obtaining an analog in the laboratory is most difficult. If we are to bring the processes of the theory into contact with the processes of behavior, however, we must either write a more elaborate theory or carry out more subtle experiments.

## Conclusions

My purpose has been to declare for an expanded consideration of human selective learning, and in the last section I suggested two directions that I think might be taken. Both were derived from an interest in the details of the processes going on between stimulus and response as they occur within any single trial. There is another direction of elaboration, that adopted by many of the stochastic learning theorists, which is to undertake an account of more of the descriptive statistics of response protocols.

There are numerous statistics available from the protocols which do not require any observations beyond the usually recorded key presses of the subjects. These statistics, some of which were mentioned previously, are generally computed for variable and sequential aspects of the data. In some cases (see R. C. Atkinson & Estes, 1963) there has been marked success in predicting the numerical values of many such statistics by using only a small number of parameter values fitted from the data.

Broadening the domain of inquiry into selective learning in this way means asking the theory to deal more exhaustively with the data collected. Broadening it in the direction of a concern for dynamic processes means adjusting the foundations of the theory to support an additional input of data instigated by an expanded experimental methodology.

One of the important sections of Noble's chapter is the survey of contemporary selective learning apparatus. He lists the major pieces of research equipment, gives references to sources of additional information, and provides a valuable summary of the characteristics of each. One of the first requirements for advances in method is a clear picture of what existing equipment and procedures accomplish for us now, and this picture has been presented in this apparatus section. Further development of new equipment, and new methods more generally, is very important to the growth of our knowledge about human selective learning. The mutual dependence between theory and experiment that typifies psychology as a laboratory science will carry this study into a broader framework, I am sure, but explicit recognition of the self-imposed constraints within which we are working may speed our rejection of those constraints.

## REFERENCES

Atkinson, J. W. *An introduction to motivation.* Princeton, N.J.: Van Nostrand, 1964.

Atkinson, R. C., & Estes, W. K. Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology.* Vol. 2. New York: Wiley, 1963, pp. 121–268.

Birch, D. Incentive value of success and instrumental approach behavior. *J. exp. Psychol.*, 1964, **68**, 131–139.

Bush, R. R., & Estes, W. K. *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univer. Press, 1959.

Dulany, D. E. Hypotheses and habits in verbal 'operant conditioning'. *J. abnorm. soc. Psychol.*, 1961, **63**, 251–263.

Dulany, D. E. The place of hypotheses and intentions: An analysis of verbal control in verbal conditioning. In C. W. Eriksen (Ed.), *Behavior and awareness—a symposium of research and interpretation.* Durham, N.C.: Duke Univer. Press, 1962. pp. 102–129.

Estes, W. K. Toward a statistical theory of learning. *Psychol. Rev.*, 1950, **57**, 94–107.

Farber, I. E. The things people say to themselves. *Amer. Psychologist*, 1963, **18**, 185–197.

La Berge, D. A recruitment theory of simple behavior. *Psychometrika*, 1962, **27**, 375–396.

Spence, K. W. The relation of response latency and speed to the intervening variables and N in S-R theory. *Psychol. Rev.*, 1954, **61**, 209–216.

Tolman, E. C. The determiners of behavior at a choice point. *Psychol. Rev.*, 1938, **45**, 1–41.

# Chapter 3

# Individual Differences

MARSHALL B. JONES
*University of Florida*

In an article in the *Psychological Review* in 1954 Edgar F. Borgatta announced to psychology and to the world a theory for the explanation of all behavior. This theory was absolutely general and capable of formulation with full mathematical rigor. He called it the General Theory of Equilibrium.

"This theory states," said Borgatta, "that for a given system composed of two or more elements, the average performance of the elements may be assessed. Then, it will be found that the performance of the *individual* elements may be specified as a direct *function* of their distance from the average performance. Further, if the direction of difference is maintained, *the sum of the differences will total to zero.* In no case will it be negative. The beauty of this theory is that *it has fitted all sets of data to which it has been applied.*"

Borgatta, of course, was spoofing. Nevertheless, the distinction upon which his theory rested makes a difference. Individual behavior can always be analyzed into typical or mean behavior and the individual's deviation from it. This difference, between mean and deviant behavior, is the root fact from which have stemmed what Cronbach (1957) called "the two disciplines of scientific psychology." From it, all other differences between experimental and differential psychology are derived.

Any experiment on motor behavior involves at least two experimental conditions. The experimental variation may be distribution of practice, knowledge of results, the order in which two tasks are practiced, or many other things; but always there must be more than one experimental condition if conclusions are to be drawn.

The purpose of an experiment is to determine the effects of the experimental conditions upon other variations which may be dependent upon them. In the analysis, therefore, the dependent data are always grouped according to the experimental conditions under which they were collected; there are always at least two sets of data. The task of analysis is to compare these sets of data and to compare them as sets. However, in order to make these comparisons the analyst must first determine which of the various behaviors which appeared under a given condition was characteristic or typical of it. This behavior is always a behavior of central tendency; almost always it is the mean behavior.

For many years psychological tests were called "experiments." The usage was incorrect because the administration of a psychological test involves a single test condition. It makes no difference how standardized or how objective the test may be; nor how many tests there are; it is not an experiment. From it nothing can be learned about the condition under which the test was administered; and from it nothing can be learned about the determinants of mean behavior on the test. In the administration of one or more tests under one test condition, it is impossible to study mean behaviors because there is nothing to study; there is nothing with which to compare them. In psychological testing deviations from the mean are all that can be studied.

By the same token, the experimentalist cannot study deviant behaviors. To him behaviors which deviate from the mean are "errors." In calling them errors he certainly intends no reflection on anyone else, because they are *his* errors. Deviations from mean behavior are caused by circumstances which the experimenter has not controlled. To the extent that they exist the experimenter has chosen to study experimental conditions which are relatively unimportant determinants of the dependent behaviors.

It must be recognized, of course, that there are many determinants which an experimenter cannot control, in the narrowest sense of this word. Since every experiment must be performed at some time, the experimenter cannot manipulate any event which took place before the experiment began. He cannot manipulate the genotype of his subjects, or their childhood experiences, or what happened to them in school. An experimenter can only manipulate present variations, and not all of them. There is, for example, little he can do about atmospheric electricity. Even those procedures about which he is most concerned cannot be carried out without some error.

For the most part, it is precisely those things which experimenters cannot control that most interest differential psychologists. Genetic variations and early experience are the essence of their concern. It is a mistake, of course, to suppose that these matters cannot be studied experimentally. In animals they are, even today. However, in an experimental area which works primarily with adult human subjects the genotypes and previous experiences of these people are the principal source of individual variation within experimental conditions.

In the study of motor learning the hiatus between experimental and differential psychology is slowly being filled; but it is still deep. The experimenters are still primarily concerned with mean differences, they manipulate conditions of learning, and their results are referable to those conditions. The differential psychologists study deviations from the mean, they manipulate conditions only on occasion, and their findings are referable to events which happened before the experiment began.

About these central differences a host of others have accumulated. Experimental method is almost always some variety of analysis of variance. Differential method is almost always correlational analysis. Experimental workers think in terms of retroactive and proactive inhibition, retention curves, and habit strengths. Students of individual differences think in terms of abilities, genetic contributions, and hypothetical factors.

Even the moods of the two disciplines are different. Differential psychology is passive, speculative, and leisurely. Its data are given; they are found, not produced. In a burst of activity the differential psychologists make their observations and analyze them; then they relax into their arm chairs. The experimentalists are active, hard-boiled, and harried. Worlds are waiting to be conquered and their would-be conquerors perform experiment after experiment in an undignified scramble to discover them.

Many people have lamented this division of psychological labor and called for a "united discipline" which can speak to Nature with a "single voice." "In both applied and general scientific work," said Cronbach (1957, p. 683) in his presidential address to the APA, "psychology requires combined, not parallel, labors from our two historic disciplines. In this common labor, they will almost certainly become one, with a common theory, a common method, and common recommendations for social betterment. In the search for interactions we will invent new treatment dimensions and discover new dimensions of the organism. We will come to realize that organism and treatment are an inseparable pair and that no psychologist can dismiss one or the other as error variance."

I do not despise prophecy, but the existing division will not be easily bridged. It may be that we already have common recommendations for social betterment; but we most certainly do not have common methods and common theories. And it is not at all clear how we might go about getting them. Perhaps, we should look for interactions, as Cronbach suggests. It may be that parallel labors will not suffice; but it may also be that the very parallelism of our labors offers the best promise of a dialogue between the two disciplines.

In motor learning mean changes occur with practice, within a single condition of learning. Since it is misleading to talk of these changes as experimental (when no experiment may have been performed), I will use the word *normative* when I wish to speak of them. In this chapter *normative* refers to mean behavior, regardless of the context in which it occurs.

Some normative events are not subject to experimentation. All human beings must walk before they can run; and it is not likely that we will soon arrange things so that running comes first; nor is there any individual variation in the order in which the two behaviors are acquired. In the sense that I use the word, walking and running are normatively ordered. Since all men must walk first and run later, walking first is also the mean behavior.

In motor learning differential changes also occur with practice. They occur at the same time, in the same place, and in the same subjects. Because they do, there exists a possibility that there may be relationships between the two kinds of events. It may be that there are correspondences between normative changes in learning and the differential process which accompanies them. There may be correspondences between the two kinds of events in the development of motor behavior or in training.

Were relationships of this sort to be found, they would not explain the results obtained in either line of inquiry. They would be empirical correspondences and no more. Nevertheless, science begins with correspondences. If we can find them in motor learning and motor behavior, we may find ourselves in need of new methods with which to study them. Someday we may even need new theory with which to explain them.

## Acquisition

Whatever its theoretical interpretation may be, learning is accomplished through trials of practice. On each trial every subject receives a score; and between scores on any two trials there exists a correlation. Since correlation is a symmetrical relationship, a correlation matrix which sets forth all these relationships contains $n(n-1)/2$ correlations, where $n$ is the number of trials. Matrices of this description play a central role in the differential study of learning.

The first correlational analysis of practice was carried out by Perl in 1934. She used four tasks and 20 trials on each task. In order to simplify the computational problem she singled out the 1st, 6th, 12th, and 19th trials for analysis. "The intercorrelations of scores on these different trials of the same task were computed," wrote Perl, "and we found that for each test they tended to be arranged in an hierarchy, i.e., the 1st trial correlated higher with the 3rd than it did with the 6th, it correlated higher with the 6th than with the 12th, and so on. The nearer together in the series trials fall the higher the intercorrelations between scores on these trials" (p. 210).

Since 1934, this finding has been repeated many times (Adams, 1953;

Edgerton and Valentine, 1935; Fleishman, 1953a; Fleishman & Parker, 1959; Greene, 1943; Houston, 1950; Melton, 1947, p. 1022; Reynolds, 1952a, b; Viteles, 1933). A study (Bilodeau, 1952) with the Two-Hand Coordination Test is a good example. The subjects were 152 basic airmen; the trials were 60 seconds long, with 30-second rest intervals between trials. The correlations between the eight trials of practice appear in Table I. Each trial correlates strongest with the trials which immediately precede or follow it. The further we move across the rows (to the right) or up the columns the greater the sequential separation and the smaller the correlation between the trials. For example, moving across the third row the correlations go: .91, .89, .87, .85, and .86. Along the fourth row they go: .91, .88, .86, and .88. In both rows there is an inversion in the last column. The last correlation in the third row should not be greater than .85; nor should the last correlation in the fourth row be greater than .86. However, with these two exceptions, and they are both minor, the matrix is organized according to the sequence in which the trials succeed one another.

This pattern is called the *superdiagonal form*. It is named after the $(n-1)$ correlations between neighboring trials, $r_{i,\ i\ +\ 1}$, which make up the superdiagonal. Each superdiagonal correlation is bottom-most in one of the vertical columns.

The superdiagonal form is an ordinal pattern. It requires only that the correlations decrease or remain the same across the rows and up the columns. The magnitude of these decrements is not specified. For example, the correlations in the 5th column go: .91, .89, .84, and .73. The third of these correlations might have been .85, .86, or .87 instead of .84, and the requirements of superdiagonal form would still have been met. There is no quantitative law which the correlations must satisfy.

TABLE I

INTERTRIAL CORRELATIONS FOR THE TWO-HAND COORDINATION TEST[a]

| *Trial* | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ |
|---|---|---|---|---|---|---|---|---|
| $x_1$ | | .79 | .77 | .74 | .73 | .71 | .71 | .70 |
| $x_2$ | | | .87 | .87 | .84 | .82 | .82 | .82 |
| $x_3$ | | | | .91 | .89 | .87 | .85 | .86 |
| $x_4$ | | | | | .91 | .88 | .86 | .88 |
| $x_5$ | | | | | | .89 | .90 | .90 |
| $x_6$ | | | | | | | .93 | .93 |
| $x_7$ | | | | | | | | .94 |
| $x_8$ | | | | | | | | |

[a] After Jones (1962).

tive positions among the individuals may be completely scrambled with no change in the standard deviation.

A correlation coefficient reflects (inversely) differences in the alignment of individuals on the two variables. The stronger the correlation the more nearly individuals are aligned alike; the weaker the correlation the greater the differences between the two arrays. A correlation, therefore, between neighboring trials of practice reflects the magnitude of the differential step between them. It measures differential change from one trial to the next.

In the matrix of Table I these (superdiagonal) correlations go: .79, .87, .91, .91, .89, .93, and .94. With one exception, the correlations increase steadily. This increase in the superdiagonal correlations means that the differential process which accompanies practice on the Two-Hand Coordination Test is negatively accelerated; as practice proceeds, the differential steps between successive trials get smaller.

Since the superdiagonal correlations are inversely related to the magnitude of the steps between neighboring trials, the process of differential change is more clearly pictured if the correlation coefficient is transformed into its inverse, the coefficient of alienation, i.e.,

$$\text{alienation} = \sqrt{1-r^2}.$$

We need also to remember that the superdiagonal correlations (or alienation coefficients) are not measures of differential status but of change in differential status, without regard to sign. They are comparable, therefore, not with mean performance but with absolute change in mean performance.

In Fig. 1 the coefficients of alienation and absolute changes from trial to trial in the mean and standard deviation of performance on the Complex Coordination Test are presented. Practice in this study (Reynolds, 1952b) was distributed into six 1-hour sessions. In each hour, performance in the first and in the last 20 minutes is plotted. There was a lunch break between the third and fourth hours.

In Fig. 1 the curve for differential alienation parallels the one for mean change. Both start high, between the first and second periods, and decrease through the third hour.

Over the lunch break there was a major change in mean performance (a drop) which was followed by generally decreasing changes (improvements) in the afternoon.

The same changes took place differentially, except that the differential effect of the lunch break was greater than its effect upon the means.

Meanwhile, the changes in standard deviation follow an irregular but level course.

In the case of the Complex Coordination Test normative and differential

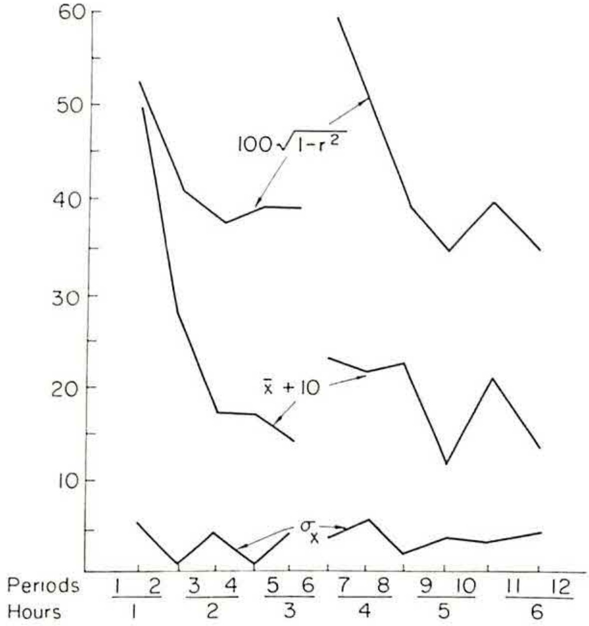


FIG. 1. Superdiagonal coefficients of alienation and changes from one trial to the next in the mean and standard deviations on the Complex Coordination Test.

changes during acquisition are roughly parallel. Nevertheless, there is in Fig. 1 a difference between the two processes which in some other tests is more marked and more sustained.

The smallest differential change during the morning hours occurred between the third and fourth periods. Between the fourth and fifth and again between the fifth and sixth periods differential changes were somewhat larger. Between the same periods changes in the mean continued to decline.

In the Rotary Pursuit (Fleishman, 1960) and Discrimination Reaction Time Tests (Fleishman & Hempel, 1955) the rate of differential change undergoes a clear change of sign. Mean change continues to decline throughout acquisition, but differential change after a pronounced initial drop increases slowly through the terminal phases of practice.

The significance of these discrepancies is that differential change may continue, in fact, it may increase, at stages of practice in which mean performance is nearing its asymptote and practice seems no longer to be having appreciable effects.

### CORRELATIONS WITH OTHER TESTS

At all stages of practice psychomotor tests correlate positively with each other. In fact, all tests of ability regardless of kind, scholastic, clerical, or sensory, printed or apparatus, subtend positive correlations. In the early days of psychological testing it was thought by some that a compensatory mechanism might be built into human nature. People who were poor at reading made up for it by being good at something else. People who had unusual intellectual gifts had poor motor coordination. However, it became apparent early in the use of psychological tests that nature was not equalitarian. People who did better than average in one area could be expected to do better than average in all other situations which called for maximal performance from the subjects. It was this fact which gave rise to Spearman's two-factor theory. Spearman hypothesized that all tests of ability depended in part upon a general factor, which he called *g*, that was responsible for the positive correlations among tests of ability.

In the mid-thirties Herbert Woodrow raised a related question. Everybody improves with practice on any test, but some people improve more than others. Woodrow wondered whether there might not be a general positivity among improvement scores in the same way that there was among raw scores on unpracticed tests. If there was, he reasoned, the fact would have important implications for the theory of intelligence.

Woodrow carried out several studies in this connection. In one of them (Woodrow, 1939a) 82 subjects were given 66 trials of practice on four tests: horizontal adding, substitution, two- and four-digit cancellation. The same subjects were also administered 21 other tests without practice. For each individual and for each of the four tests which were practiced, Woodrow calculated an improvement score, which equalled the individual's score at the end of practice minus his score at the beginning. He found that the correlations among improvement scores were low, on the order of .20 to .30. The correlations, moreover, between improvement scores and the other tests were also low. As often as not they were negative. He concluded that there was no general factor common to improvement scores.

In the same study Woodrow also reported the correlations between the 21 unpracticed tests and initial and final performance on each of the practiced tests. The results appear in Table III. In the case of all four tests, initial performance tended to correlate better with the unpracticed tests than did final performance.

The examination of psychomotor performance in relation to external tests was deferred for 20 years. In the 1950's, however, a series of studies from the Air Force confirmed and extended Woodrow's earlier results with printed tests. The principal study in this series was carried out by Adams in 1953.

The subjects in Adams' study were 197 basic airmen. They were administered 32 printed tests, 13 simple psychomotor tests, and 7 complex psychomotor tests. All of the complex psychomotor tests were practiced and

TABLE III

Average Correlations of 21 Unpracticed Tests with Initial and Final Performance on Four Practiced Tests[a]

| Test | Average correlation | |
|---|---|---|
| | *Initial* | *Final* |
| Horizontal adding | .30 | .19 |
| Substitution | .22 | .13 |
| Two-digit cancellation | .15 | .10 |
| Four-digit cancellation | .18 | .16 |

[a] After Woodrow (1939a).

none of the others. Practice was most extensive on the Complex Coordination Test (CCT), which served as the focal point of the investigation.

Adams' results are presented in Table IV. On the average, the printed tests correlated better with initial performance than with final performance on the CCT. The 13 simple psychomotor tests maintained their relationship to initial performance on the CCT, even increased it a little. Initial performance on the six complex psychomotor tests other than the CCT behaved like the printed tests; it correlated better with initial than with final performance on the criterion test. However, final performance on the same six tests correlated better with final performance on the CCT than with initial performance.

These results have been repeated by several investigators using many different psychomotor tests. The results have uniformly confirmed Adams'

TABLE IV

Average Correlations of 32 Printed, 13 Simple Psychomotor, and 6 Complex Psychomotor Tests with Initial and Final Performance on the Complex Coordination Test[a]

| Kind of test | Average correlation with the CCT | |
|---|---|---|
| | *Initial* | *Final* |
| Printed | .38 | .25 |
| Simple psychomotor | .24 | .27 |
| Complex psychomotor | | |
| Initial | .45 | .38 |
| Final | .41 | .49 |

[a] After Adams (1953).

findings. If two tests are widely dissimilar, practice on one of them generally serves to reduce the correlation between the two tests. If they are alike, practice tends to increase the relationship between the tests.

## THEORY

In a general way, there is considerable agreement as to the interpretation of the differential process which accompanies practice. Students of the subject are agreed that the abilities at work in successive trials become progressively fewer with practice. The hypothesis was first proposed by Bradley Reynolds in 1952. Fleishman and Hempel (1953) reached the same conclusion through factor analysis; and I reached it on correlational grounds but without factor analysis (Jones, 1959, 1960). There is, however, disagreement as to the formal expression this consensus should take. Two theories dominate: a simplicial hypothesis and a factor-analytic one.

In the case of an eight-trial matrix the simplicial hypothesis posits that

$$\begin{aligned}
x_1 &= f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 + s_1\\
x_2 &= f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 + s_2\\
x_3 &= f_1 + f_2 + f_3 + f_4 + f_5 + f_6 \phantom{{}+f_7} + s_3\\
x_4 &= f_1 + f_2 + f_3 + f_4 + f_5 \phantom{{}+f_6+f_7} + s_4\\
x_5 &= f_1 + f_2 + f_3 + f_4 \phantom{{}+f_5+f_6+f_7} + s_5\\
x_6 &= f_1 + f_2 + f_3 \phantom{{}+f_4+f_5+f_6+f_7} + s_6\\
x_7 &= f_1 + f_2 \phantom{{}+f_3+f_4+f_5+f_6+f_7} + s_7\\
x_8 &= f_1 \phantom{{}+f_2+f_3+f_4+f_5+f_6+f_7} + s_8,
\end{aligned}$$

where $x_i$ is an individual's score on the $i$th trial and $f_m$ is his score on the $m$th factor. The $f_m$ are common factors, i.e., they are operative in at least two trials. The factors $s_i$ are specific factors which operate in a single trial only.

The simplicial hypothesis states the consensus directly: $f_7$ is present in the first two trials and then drops out; $f_6$ acts in the first three trials only; $f_1$ is operative in all eight trials.

These factors should not be understood as unitary; they are composites of all differential elements which act in the first two trials, the first three, all eight, or whatever the span that the factor covers may be.

The most powerful argument in favor of the simplicial hypothesis, aside from the clarity with which it states the consensus, is that the law of single tetrad differences follows from it. Under a simplicial hypothesis the correlation between trials $i$ and $j$ is

$$r_{ij} = \left(\frac{\Sigma_i}{\Sigma_j}\right)^{1/2} \left(\frac{\Sigma_i}{\Sigma_i + \sigma^2_{s_i}}\right)^{1/2} \left(\frac{\Sigma_j}{\Sigma_j + \sigma^2_{s_j}}\right)^{1/2},$$

where

$$\Sigma_i = \sigma^2_{f1} + \sigma^2_{f2} + \ldots + \sigma^2_{fi}$$

Therefore, the tetrad equality

$$r_{ik} \times r_{jl} = r_{il} \times r_{jk}$$

reduces to

$$\left(\frac{\Sigma_i \Sigma_j}{\Sigma_k \Sigma_l}\right)^{1/2} = \left(\frac{\Sigma_i \Sigma_j}{\Sigma_l \Sigma_k}\right)^{1/2},$$

which is an identity.

In addition to explaining the intertrial correlations and the law which governs them, the simplicial hypothesis also explains the relationships between trials of practice and other tests. The factors which are present when practice begins but then drop out are more general than those which remain; they have a broader span. Tests, therefore, which are dissimilar to the one which is practiced are likely to relate to it through these more general factors. When they drop out, the intertest correlation goes down.

If the two tests are alike, the relationship between them is likely to rest on those factors which persist in practice. Hence, the intertest correlation does not decrease with practice; and it may increase. The correlation between initial performance on two psychomotor tests is bolstered by widespread factors which subsequently drop out. At the same time, however, those factors which remain comprise less of the total variance in the early trials than they do later on. Therefore, if persistent factors in the two tests are closely related, the gain in intertest covariance which comes about through their increasing prominence may more than compensate for the loss of covariance from the dropping out of more widespread factors.

In the idea that differential elements become fewer with practice, we have a parallel with normative changes. Since the days of Bekhterev, experimental psychologists have been agreed that irrelevant responses are eliminated with practice. When they received shock for the first time, Bekhterev's dogs jumped with all four feet. At the end of training they lifted their right front paw just the amount necessary to avoid shock. Similar processes take place in human learning. Practice, whether it is viewed normatively or differentially, seems to be a process of elimination, both of responses and of differential elements.

The factor-analytic hypothesis is more difficult to state. In her original study of trial intercorrelations, Perl (1934) performed a factor analysis. She says of the results that, "in each case the factor loadings seem to increase from the first trial to about the middle trial used, and then decrease again to the end of the series. The second factor loadings seem to form an hierarchy going from a fairly large plus on the first trial to a low plus, then to a low minus and then to a larger minus on the last trial or else the same thing with the signs reversed" (p. 210). A year later Edgerton and Valentine (1935) repeated Perl's study with different tasks; they obtained the same results that she had, with respect both to pattern of correlation and the results of factor-analyzing it.

Though these early investigators did not realize it, there was an analytic connection between their empirical and factor-analytic results. The first person to appreciate this connection was Herbert Woodrow. "These investigators," wrote Woodrow (1939b, p. 459), "obtained a result which could have been predicted from the fact that the trials following the first show less and less correlation with the first trial and more and more correlation with the last trial. The factor analysis of such a series is bound to give at least two factors. . . ." Then Woodrow described the results that Perl and Edgerton and Valentine had obtained.

Woodrow's statement of this connection, though correct, was incomplete. It remained for Louis Guttman in 1954 to state it fully. The principal components of any matrix which has the superdiagonal form obey what Guttman called the "law of oscillations." The loadings on the first component are all positive and bigger in the middle than at either end. The loadings on the second component undergo one change of sign; either they begin positive and become negative or vice versa. The loadings on the third component undergo two changes of sign; if they are plotted against trials, the locus is U-shaped. The loadings on each additional component undergo an additional change of sign.

Guttman's law is one form which the factor-analytic hypothesis may take. As an interpretation of the superdiagonal forms of practice, it is open to two major objections. The first is that it does not state the consensus that practice is a process of elimination. The law of oscillations presupposes that the variables which enter into the matrix have been arranged in serial order; in this respect it differs from the usual factor-analytic formulation, which takes no account of order. However, the pattern which it requires of the factor loadings, though orderly, stands in no obvious relation to simplification.

A consequence of this failure is that the factor-analytic hypothesis does not explain the relationships between trials of practice and other tests. It is the idea of simplification which makes possible an understanding of these relationships, and the factor-analytic hypothesis does not contain it. In principle, the decrease with practice in the correlation between printed and psychomotor tests might be explained in some other way. But the law of oscillations gives us no inkling as to what it might be.

The second objection is that the law of single tetrad differences does not follow from the hypothesis. It is true that all superdiagonal forms generate factor structures which obey Guttman's law; but it is not true that all factor structures which obey the law of oscillations generate superdiagonal forms. The direction in which the logic runs is just the reverse of what it ought to be. The "theory" follows from the data, instead of the other way around.

Another way to put the factor-analytic view is to list the loadings of every trial on every factor. In this form the hypothesis explains the correlations. If each trial loads on hypothetical factors in the manner specified by the hypothesis, then each correlation ought to be what it is.

This second formulation, however, is subject to the same two objections as the first. It does not state the consensus but presents us with lists of numbers. And the law of single tetrad differences does not follow from it. In this second form, the factor-analytic hypothesis is numerical. It reproduces the correlations; but it does so by enumeration, without explaining the law which rules them.

## EXPERIMENTAL STUDIES

Differential processes take place during practice with no change in controlled conditions. If these conditions are changed, then the form and content of the differential process may change with them. Fleishman (1957), for example, has shown that if the same task is used but its difficulty is changed, then the correlations of the trials with reference tests change also.

Task difficulty, of course, is only one among many experimental variations that may be studied. In another study Fleishman (1953a) gave standard and experimental forms of the Rudder Control Test to two groups of air cadets. One group took the standard and the other the experimental test first. Since there were six trials of practice on the standard and four on the experimental test, there were 15 correlations within the standard test, 6 within the experimental, and 24 between the tests, under either order of administration. In Table V the averages for each of these groups of correlations are presented.

With respect to differential relationships the standard test was more disturbed by following the experimental test than vice versa.

For both tests the average within-test correlation was lower when the test came second than when it came first; but the drop was greater for the standard test.

The average correlation between tests was greater when the standard test came first than when it came second.

The patterns of within-test correlations are not presented, but they were superdiagonal forms for both tests under both administrations. However, the pattern was rougher for the standard test when it came second than when it

TABLE V

AVERAGE CORRELATIONS WITHIN AND BETWEEN STANDARD AND EXPERIMENTAL FORMS OF THE RUDDER CONTROL TEST WHEN THE STANDARD FORM WAS ADMINISTERED FIRST AND WHEN THE EXPERIMENTAL FORM WAS ADMINISTERED FIRST[a]

| *Correlation* | *Standard first* | *Experimental first* |
|---|---|---|
| Standard-standard | .64 | .48 |
| Experimental-experimental | .76 | .83 |
| Standard-experimental | .43 | .32 |

[a] After Fleishman (1953a).

came first, while the patterns of correlation within the experimental test were smooth superdiagonal forms under both administrations.

When the standard test came first, the pattern of between-test correlations was very regular; in fact with the standard test first, the *whole* matrix had regular superdiagonal form. However, with the experimental test first, the pattern of between-test correlations was rough.

It is significant that in neither of these studies did Fleishman report any normative results. The usual thing is the opposite; normative results are reported but not differential. Nevertheless, in any experimental study of motor learning both kinds of information are available. We know a great deal about the normative effects of many experimental variations; but we know next to nothing about the differential consequences of the same variations. And there may be relationships between the two orders of events. It may be, for example, that the amount of differential disturbance produced in one task when it follows upon another bears a relationship to transfer of training between the two tasks.

## Relearning and Recall

The usual approach to the study of memory is through retention or relearning. A task is learned and then, at varying intervals afterwards, one or several additional trials are administered. The retention or relearning trials do not differ in their instructions or procedures from the trials of original learning. The subject is not asked to recall how or what he did in original learning; he is asked again to perform the task as best he can.

Studies of motor skill retention have uniformly found curves which decline at a slow rate. And the losses are quickly regained in short periods of relearning. In only one study (Fleishman & Parker, 1962) have correlational results been reported; it concerned a complex tracking task. Three groups of subjects (of 10 each) learned the tracking task "on their own," without additional instruction. Each group was tested for retention once, after 9, 14, or 24 months of no practice. Four more groups of subjects (also of 10 each) were trained with supplementary demonstration and explanation of the task. Each of these groups was also retested once, after 1, 5, 9, and 14 months.

In all seven groups there was very little mean decrement in performance over periods of no practice as long as two years. Original learning reached slightly higher levels, for the same amount of practice, in the groups which had received supplementary instruction; and retention was slightly higher in these four groups.

Correlations between individual performance at the end of original training and in retention testing were high, ranging from .80 to .98; they were slightly higher for the groups which had received formal instruction. Sixty-four percent of the variance in individual performance at retesting was attributable to differences in original learning and only 10% to the retention interval.

In order to avoid practice effects, experimental studies of psychomotor retention have used different groups of subjects to determine retention over different periods of time. The consequence is that we do not have correlations among retention trials, but only between each retention trial and original learning. However, retention trials do not differ from trials of practice except for the long intervals of time between them; and correlations between trials of practice take the superdiagonal form regardless of massing or distribution of practice. We have every reason, therefore, to expect that correlations among retention trials would also have the superdiagonal form.

An alternative approach to the study of memory is to ask a subject not to do the best he can on a previously practiced task but to recall a motor or verbal event which took place at some time past. In the past few years E. A. Bilodeau and his co-workers have conducted several experiments in which recall, i.e., remembrance of past events, rather than relearning or retention has been studied.

In one of these studies (Bilodeau, Jones, & Levy, 1964) subjects were first asked to move a manual lever through an arc until it ran up against a stop. The stop was removed and the subject asked to move the lever to the point where the stop had been. This response was called $R_0$. The subject was told that he had missed the target point by an arbitrary number of units. This information was called $KR_0$. After 3 minutes, 20 minutes, 2 days, and 6 weeks the same subjects were asked to recall $R_0$ (by moving the lever) and

$KR_0$ (verbally). These responses were called $R_1,...,R_4$ and $KR_1,...,KR_4$, respectively.[1]

In Table VI the correlations among $KR_1,...,KR_4$ are presented. They have the familiar superdiagonal form. Neighboring attempts at recall correlate more strongly with each other than more separated attempts.

The correlations among $R_1,...,R_4$ appear in Table VII; and they do not have the superdiagonal form. Instead, the correlations fall into a "monotonic hierarchy." The largest correlation is in the upper left-hand corner, the smallest is in the lower right-hand corner, and in between the correlations decrease as we move across the rows (to the right) or *down* the columns.

Monotonic hierarchies, like superdiagonal forms, are ordinal patterns. They do not necessarily obey any exact regularity; but they may, as in this instance they do. The correlations in Table VII obey Spearman's law of *double* tetrad differences. The two differences equal .001 and .027.

TABLE VI
CORRELATIONS AMONG FOUR SUCCESSIVE ATTEMPTS TO RECALL $KR_0$ IN THE EXPERIMENT BY BILODEAU, JONES, AND LEVY (1964)

| *Recall Attempt* | $KR_1$ | $KR_2$ | $KR_3$ | $KR_4$ |
|---|---|---|---|---|
| $KR_1$ | | .87 | .92 | .39 |
| $KR_2$ | | | .94 | .39 |
| $KR_3$ | | | | .49 |
| $KR_4$ | | | | |

TABLE VII
CORRELATIONS AMONG FOUR SUCCESSIVE ATTEMPTS TO RECALL THE LEVERING RESPONSE $R_0$ IN THE EXPERIMENT BY BILODEAU *et al.* (1964)

| *Recall attempt* | $R_1$ | $R_2$ | $R_3$ | $R_4$ |
|---|---|---|---|---|
| $R_1$ | | .72 | .61 | .44 |
| $R_2$ | | | .58 | .48 |
| $R_3$ | | | | .38 |
| $R_4$ | | | | |

---

[1] The experiment was more complicated than it appears in this account. There was another group of subjects and an additional response, neither of which is mentioned here.

This pattern is exceptional in correlations among successive trials which are normatively the same. The exception involves recall of a motor event. How significant this circumstance is remains to be determined. In any event, this particular matrix cannot be interpreted in the same terms as the superdiagonal forms of practice or recall of KR.

If we suppose that successive attempts to recall the levering response involve a single common factor, i.e., if

$$\begin{aligned} \mathrm{R}_1 &= g + s_1, \\ \mathrm{R}_2 &= g + s_2, \\ \mathrm{R}_3 &= g + s_3, \\ \mathrm{R}_4 &= g + s_4, \end{aligned}$$

where $s_i$ is specific to $\mathrm{R}_i$, the correlations will form a monotonic hierarchy which obeys the law of double tetrad differences.

In their interpretation of these results Bilodeau *et al.* identified $g$ with the memory trace of $\mathrm{R}_1$.[2] With the passage of time and the intervention of retroacting events (including previous attempts at recall), this trace is attenuated by increasing amounts of random error, the $s_i$. The correlations of each recall attempt with the one common factor,

$$r_{\mathrm{R}_i g} = \frac{\sigma_g}{\sigma_{\mathrm{R}_i}}.$$

In the case of Table VII these correlations are .85, .85, .70, and .55 for $\mathrm{R}_1$, $\mathrm{R}_2$, $\mathrm{R}_3$, and $\mathrm{R}_4$, respectively. As random error accumulates, the accuracy of recall becomes progressively worse.

The variances of successive recall attempts increase with time and in a manner roughly proportional to the decay of $r_{\mathrm{R}_i g}$. What appears, therefore, to happen is that each successive attempt at recall is attended by an increment in unrelated variance which obfuscates retrieval of $\mathrm{R}_1$.

The sources of this accumulating specific variance are not clear from the experiment by Bilodeau and his co-workers; nor is the prevalence of monotonic hierarchies in the recall of motor and verbal events. But the occurrence of a pattern in successive trials, whether of learning, retention, relearning, or recall, which is not superdiagonal is at least refreshing; and it may be more.

---

[2] The correlations for $\mathrm{R}_0$, the training response, though they had the correct form, i.e., decreasing with time, did not participate in the monotonic hierarchy; they were too low. For some reason, the subjects apparently chose to model their later responses of recall not on the training response as demanded by the experimenter but on the first recall response. Only one training trial was used and the subjects were not told that they would be tested for retention; these procedures may have contributed to the low level of correlation for $\mathrm{R}_0$.

## Inheritance

If there is any one book with which differential psychology may be said to have begun, it is Francis Galton's *Hereditary Genius*. The bulk of Galton's text was concerned with the inheritance of those intellectual and personal qualities which make for eminence in judges, commanders, literary men, musicians, scientists, painters. However, Galton also considered familial tendencies in two distinctly motor vocations, rowing and wrestling.

"In respect to oarsmen," said Galton (1887, p. 305), "let me assure the reader that they are no insignificant fraction of the community, – no mere waifs and strays from those who follow more civilized pursuits. A perfect passion for rowing pervades large classes. At Newcastle, when a great race takes place, all business is at a standstill, factories are closed, shops are shut, and offices deserted. The number of men who fall within the attraction of the career is very great; and there can be no doubt that a large proportion of those who are qualified to succeed brilliantly, obey the attraction and pursue it."

Galton compiled information on all men who had rowed with distinction on either the Tyne or the Thames in the preceding 26 years. "The names," said Galton, "are not picked and chosen, but the best men have been taken of whom any certain knowledge could be obtained." Galton classified his oarsmen into two broad categories of eminence: good and excellent. He calculated that not 1 man in 10 attained the lower rank and less than 1 in 1,000 became an excellent oarsman.

Galton found a total of 35 men who had attained a degree of eminence as oarsmen sufficient to rank them in either of these two grades. Of these 35 men 33 were related to one or more of the other eminent oarsmen. There were 6 families with 2 eminent oarsmen, 2 with 3, 1 with 5, and 1 with 10. The most remarkable of these was the family of Clasper; the family relationships of the ten oarsmen appear in Fig. 2.

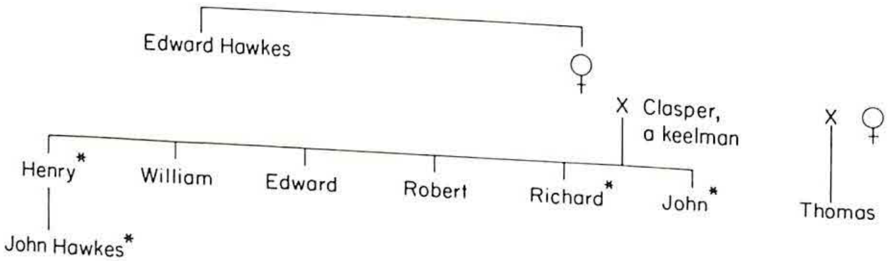


FIG. 2. The family of Clasper (after Galton, 1887). Names marked with an asterisk are excellent oarsmen; others are good.

The eponym of the family was Henry Clasper. "He was," wrote Galton, "for many years stroke of a four-oared crew, and frequently the whole crew, including the coxswain, were members of the Clasper family. For eight years this crew won the championship of the Tyne. Six times Henry pulled stroke for the crew winning the championship of the Thames, and Coombes declared that he was the best stroke that ever pulled." Henry's brother, Richard, was only slightly less distinguished, and another brother, Edward, ranked as a good rower despite the disadvantage of having lost a leg.

Galton's studies of inheritance have been criticized principally on the grounds that familial tendency is not conclusive evidence of hereditary gifts. In a family like Clasper's rowing must certainly have dominated the ambitions of all its male members. At a very early age they must have prepared for an oarsman's career. And this preparation may have had a bearing on their phenomenal success.

The figures on oarsmen may also be questioned on technical grounds. According to Galton's estimates, the ratio of good to excellent oarsmen was roughly 100 to 1. Nevertheless, in the list of names he gave the ratio was 1 to 1; there were as many excellent as good oarsmen. The most likely reason for this discrepancy is that it was more difficult to obtain "certain knowledge" about less distinguished men. Still, if the 1600-odd good oarsmen which Galton was not able to locate had also been listed, it is possible that Galton's figures would have been less impressive.

With human subjects it is not easy to distinguish familial from genetic transmission. An approach which reduces the extent of the problem is to study young children. Individual differences in behavior are present in the first few months of life and persist without much change through the first two years, at least (Thomas, Chess, Birch, Hertzig, & Korn, 1964). Some of these differences, like activity or the responsiveness of a child to external stimulation, involve heavy motor components.

The study of young twins carries the argument further. Freedman and Keller (1963) followed 11 pairs of fraternal and 9 pairs of identical twins on a monthly basis through the first year of life. The twins' behavior was evaluated on the Bayley Mental and Motor Scales and the Bayley Infant Behavior Profile. Within-pair differences in all scales were much smaller between identical than between fraternal twins.

In twin studies environmentalists usually advance the argument that identical twins tend to imitate each other, to identify with each other, and that their observed likeness to each other is secondary to these psychological processes rather than their genetic identity. However, in Freedman and Keller's study this argument does not apply. In the first year of life infants do not respond to each other at all, much less imitate each other.

Similar differences between identical and fraternal twins have also been observed at the adult level. In 1933 McNemar studied the performance of 46 pairs of fraternal and 47 pairs of identical twins on the Koerth pursuit rotor, steadiness, speed drill, spool packing, and card sorting tests. Holzinger heritability coefficients appear in Table VIII; these coefficients state the proportion of the variance within families which is attributable to genetic variations. For the pursuit rotor, steadiness, and speed drill tests the heritabilities are high; for card sorting heritability is moderate and for spool packing, it is low.

On three of these tests, pursuit rotor, spool packing, and card sorting, the twins were given seven trials of practice. In Table IX Holzinger heritabilities for the first two and the last two trials are presented. In all three cases these heritabilities tend to be lower at the end of practice than at the beginning. In the case of the pursuit rotor and card sorting tests the decline is relatively slight and easy. In the spool packing test, there is no heritable variance except on the first trial; thereafter, it disappears altogether.

TABLE VIII
HOLZINGER HERITABILITY COEFFICIENTS (CORRECTED FOR AGE) IN 46 PAIRS OF FRATERNAL AND 47 PAIRS OF IDENTICAL TWINS ON FIVE PSYCHOMOTOR TESTS[a]

| *Test* | *Heritability* |
|---|---|
| Pursuit rotor | .90 |
| Steadiness | .80 |
| Speed drill | .69 |
| Spool packing | .25 |
| Card sorting | .46 |

[a] After McNemar (1933).

TABLE IX
HOLZINGER HERITABILITY COEFFICIENTS (CORRECTED FOR AGE) IN 46 PAIRS OF FRATERNAL AND 47 PAIRS OF IDENTICAL TWINS ON TRIALS 1, 2, 6, AND 7 IN THREE TESTS OF MOTOR SKILL[a]

| *Test* | 1 | 2 | 6 | 7 |
|---|---|---|---|---|
| Pursuit rotor | .78 | .74 | .64 | .67 |
| Spool packing | .28 | −.02 | −.17 | −.02 |
| Card sorting | .43 | .51 | .31 | .43 |

[a] After McNemar (1933).

There is one other study in the literature which concerns itself with the magnitude of genetic contributions at different stages of practice (Brody, 1937); it found no change with practice on the Minnesota Spatial Relations Test.

It appears, therefore, that genetic contributions tend to be more prominent at the beginning of practice than later on. However, these twin studies are now almost 30 years old; it might be profitable to take up these questions again.

## Task Classification

Many attempts to group psychomotor tests into categories, to develop some sort of task classification, are normative. They are based on features of the task which do not vary with the individuals who work at it. The trichotomy of printed, manipulative, and apparatus tests is a normative breakdown. So is the contrast of short and long tests. In standardized tests, materials and procedures do not vary from one person to the next.

A second kind of taxonomy is differential. In a general way, two tests fall into the same category if they are strongly correlated with one another, and into different categories if they are not.

Between these two kinds of taxonomy there exists an odd sort of correspondence. In a group of tests which correlate positively with each other it is often possible to indicate the normative elements upon which the differential communality appears to rest. It may be that all of the tests involve rapid movements of the wrists and fingers or eye-hand coordination or different sense modalities. In other cases, the normative communalities are unclear or only guessed at.

The correlation between two tests is a precise statement of their differential similarity. Any two correlations may be compared directly, because they are absolute numbers, entirely independent in their terms of the tests which subtend them. For this reason, the dominant taxonomies in psychomotor learning as everywhere in testing are differential. Unfortunately, the uncertainty of their normative bases is equally general.

This section of the chapter is entitled "Task Classification," which, though it may seem innocent, is not. It reflects a bias on my part which is not shared by the majority of my colleagues. Most differential psychologists would have called this section "Psychomotor Abilities." This difference in language recalls a tradition which goes back to instinct and faculty psychology. According to the majority view, psychomotor learning is a "domain" of behavior. Individual differences within it are classified into a finite number of "human abilities." And the task of the differential psychologist is to "measure" these abilities. If he succeeds in this effort, then the whole domain of psychomotor learning has been properly described and accounted for.

The alternative view is that the number of abilities in psychomotor learning is indefinitely large. It points out that the number of abilities which psychologists have found necessary to describe the various domains of behavior has steadily increased. There was a day when the major question which troubled students of these matters was whether or not all of the aptitude domains were to be understood in terms of one or several abilities. For a while in the 1930's psychologists felt that they had identified seven so-called primary abilities. Twenty years later (1956) Guilford found it necessary to posit some 40 or so abilities in the cognitive domain alone.

It also calls attention to the fact that even the factor analysts name their abilities after the tests in terms of which they are identified. Individual differences on a single task are referable to the various biographies of the subjects. But the relationship between two tests is a function of the tests.

According to the minority view, differential elements arise in the first instance and are organized according to genetic, physiological, and learning principles which bear no essential relationship to the correlations we observe among tests. These correlations are determined by the tests, by the specific combinations of elements upon which they draw. Correlations among tests reflect the organization of the tests, not the people who take them.

If this view is correct, the number of factors which can be discovered in any area is not limited by any organization inherent in human beings. It is limited only by the industry and creativity of test makers.

## Differential Factors in Psychomotor Tests

Existing classifications of differential components in psychomotor tests are exclusively factor analytic. In this technique the variance of each test is resolved into the test's loadings on several hypothetical factors; the resolution is effected through analysis of the correlations among the tests.

There has always been some doubt as to what it is that factor-analytic resolutions assert about the correlations from which they are derived; but this chapter is an inappropriate place in which to attempt a clarification of the question. However the matter is put, factor analysts have reached agreement with respect to several factors which operate in psychomotor tasks (Fleishman, 1953b, 1954).

A factor called *wrist-finger speed* has been identified in several tapping and aiming tests. Its loadings suggest that it is independent of eye-hand coordination. For example, the speed with which a person can tap a pencil, without regard to where he taps it, loads high on the factor. However, if taps are scored only when they fall within a small circle, tapping speed loads low; and the smaller the circle the lower the loading.

*Aiming* is supplementary to *wrist-finger speed.* It also involves speed, but tests which require precise positioning of each tap load highest on the factor. Eye-hand coordination appears to be essential. The factor appears in printed tests only.

The Purdue Pegboard and O'Connor Finger Dexterity Tests have high loadings on *finger dexterity*. Both tests require precise manipulation of small objects, like pegs or pins. Sensing with the fingertips is essential.

*Manual dexterity* is like *finger dexterity* except that it concerns larger objects. Tests which load on it do not require sensing with the fingertips.

Tests which require a subject to hold an object stationary, for example, to hold a stylus so that it does not touch the sides of a small hole, load heavily on *steadiness*. The factor also appears when subjects are asked to move an object in a highly controlled manner, for example, in a track so that the stylus does not touch the sides.

*Psychomotor coordination* is found almost exclusively in apparatus tests, like the Rotary Pursuit, Rudder Control, and Complex Coordination Tests. The normative basis of this factor has been rather vaguely described as "representing either coordination of the body, in movements of moderate scope, or coordination of such movements with the perception of a visual stimulus."

Other factors in psychomotor tasks have been described, but the six which have been mentioned are among the most frequently found.

### INTERNAL STRUCTURE

Preoccupation with the relationships between psychomotor tasks has tended to obscure the equally important question of their internal structure. If a task consists of several identifiable components, how does performance on any one or more of these components relate to performance on the whole task? There are very few studies in the literature which deal with this question.

In one of them, E. A. Bilodeau (1957) used a single subject. This man, who "had received a very considerable amount of practice and had demonstrated a stable level of performance," was administered 48 trials on each of four tasks on the Multidimensional Pursuit Test. The most inclusive of these tasks contained four components. These components corresponded to meters, one each for heading (H), altitude (A), bank (B), and air speed (S). The subject's task was to keep all of the meters centered through manipulation of appropriate controls.

The subject's performance was evaluated as the proportion of each 1-

minute trial that he was on target. There were, therefore, four time-on-target scores in each trial, one for each meter.

There were also six further time-on-target scores for the meters taken in pairs. For each pair of meters the score was the proportion of the time that he was on target in *both* meters.

Similarly there were four time-on-target scores for the meters taken in threes. The score was the proportion of the time that the subject was on target in all three meters.

And finally there was the proportion of each trial that the subject was on target in all four meters.

Bilodeau found that a simple multiplicative rule controlled the relationships between time-on-target scores for individual meters and composite scores for two or more meters. In any given trial, if the subject was on target half of the time in meter A and half of the time in meter B, he was on target in both of them one quarter of the time. If his time-on-target scores for the four meters individually were

$$\begin{aligned} A &= .8, \qquad & B &= .5, \\ H &= .7, \qquad & S &= .6, \end{aligned}$$

then his time-on-target scores in the same trial for the composite scores were

$$\begin{aligned} (AB) &= (A)\,(B) = .8 \times .5 = .40, \\ (ABH) &= (A)\,(B)\,(H) = .8 \times .5 \times .7 = .28, \\ (ABHS) &= (A)\,(B)\,(H)\,(S) = .8 \times .5 \times .7 \times .6 = .168, \end{aligned}$$

and similarly for the other combinations. Calculating predicted scores according to this rule Bilodeau obtained correlations over trials between predicted and observed values of .89, .88, and .91 for the composites (AB), (ABH), and (ABHS), respectively.

Bilodeau also found that the correlations among time-on-target scores for the individual meters were low. It followed from this result that the same rule which governed the trial scores also held for the means over trials, i.e., that

$$\begin{aligned} (\overline{AB}) &= (\bar{A})\,(\bar{B}), \\ (\overline{ABH}) &= (\bar{A})\,(\bar{B})\,(\bar{H}), \\ (\overline{ABHS}) &= (\bar{A})\,(\bar{B})\,(\bar{H})\,(\bar{S}). \end{aligned}$$

For this one subject, average performance on the composite tasks was a simple multiplicative function of average performance on each of the component tasks.

This study concerned intratask relationships in one individual. It was not, therefore, directly concerned with individual differences. However, in another study Bilodeau (1955) administered the Multidimensional Pursuit

Test to 72 basic airmen. He found that the same multiplicative rule which held within a subject also held between subjects. He also found that individual differences in performance on the single meters, A, B, H, and S, were roughly independent of each other.

In this 72-subject study Bilodeau did not present correlations among the various components. But it wasn't necessary. If the components were arranged in order of increasing complexity, i.e., A, AB, ABH, ABHS, the correlations would necessarily fall into a superdiagonal form which obeyed the law of single tetrad differences.

This same conclusion also applies to correlations among A, AB, ABH, and ABHS, within a subject. In his single-subject study Bilodeau presented these correlations; they appear in Table X. The correlations have the superdiagonal form and the tetrad difference equals .022.

This result is general. In a series of tasks such that each member in the series includes the components of every task which precedes it, the correlations between tasks will fall into the superdiagonal form and obey the law of single tetrad differences. Only one hypothesis is necessary, namely, that individual differences in the component behaviors are independent. It makes no difference whether the more complex tasks are additive or multiplicative functions of their parts. Even when more complicated functions are considered, it matters very little what they are; as long as the component behaviors are differentially independent, the correlation pattern follows from the inclusion of each task in those which follow it.

In these propositions we have a fundamental connection between normative and differential relationships. The parts of a task are defined by its materials and procedures. Though different individuals perform differently in all of them, the parts themselves are normatively defined. In the Multidimensional Pursuit Test every subject performed in some fashion with respect to all four meters; the composition of the task was the same for everyone.

TABLE X

CORRELATIONS AMONG FOUR TASKS OF INCREASING COMPLEXITY IN A SINGLE SUBJECT[a]

| *Task* | *A* | *AB* | *ABH* | *ABHS* |
|---|---|---|---|---|
| A | | .73 | .58 | .51 |
| AB | | | .84 | .70 |
| ABH | | | | .88 |
| ABHS | | | | |

[a] After Bilodeau (1957).

Superdiagonal patterns do not necessarily imply normative inclusion. The superdiagonal forms of practice are obvious examples. In practice the normative requirements of the task do not change. What happens is that each successive trial draws upon fewer and fewer elements in the differential background of the subjects.

(It is also worth noting that in practice complexity decreases with time, whereas in the building of tasks one upon another complexity typically increases with time. The superdiagonal form is indifferent with respect to the relationship of time to increasing or decreasing complexity.)

Nevertheless, the basic principle remains: wherever we observe that tasks are built one upon another, we know that the correlations among them will take the superdiagonal form as long as the tasks are composed of parts which are differentially independent.

## Selection

Personnel selection is the principal point of application of differential psychology. In any training or employment program some people do better than others, and some attrite. If there are enough applicants and their prospects can be identified in advance, admissions to the program can be limited to applicants who are likely to do well.

The procedure for approaching this problem has become standard. First, a selection battery is constructed. Biographical information is collected from a sample of applicants and a variety of tests is given. Among those applicants who are admitted to training or employment, correlations are computed between criterion performance, proficiency or attrition, and each individual test or item. The tests and items are then combined into an additive composite which correlates maximally with the criterion.

Once constructed, the selection battery is cross-validated. The tests or items which enter into the battery are administered to a new sample of applicants and the correlation is computed between the battery composite, as determined in the first sample, and the criterion. This second correlation is the validity coefficient of the battery.

In practice this procedure is rarely carried out in detail. A test which was found useful in an allied program is simply taken over and used. Frequently no check is made on its validity, even in large governmental programs. Sometimes, as in the OSS program during World War II, little or no thought is given to what the criterion is supposed to be.

Almost all psychological tests used in personnel selection are printed. In programs in which the criterion behaviors may be assessed by printed tests, for example, in school situations, predictive validities occasionally range above .50, but not often. In programs in which the criterion behaviors cannot be assessed by printed tests, as in pilot training, dentistry, machine operation, or any other genuinely psychomotor vocation, predictive validities in excess of .50 are unknown; usually they are much less. The Navy pilot selection program has always relied exclusively upon printed tests. Predictive validities in the Navy program have rarely reached the .40's; in most samples they fall in the .30's.

It needs to be recalled that these coefficients are based on *admitted* applicants. If the figures were corrected so as to estimate what the correlations would be if all applicants were admitted, validities would be higher.

Simple manipulative tests, like the O'Connor Finger Dexterity, Purdue Pegboard, and Minnesota Rate of Manipulation tests, have been widely used in industry. The O'Connor test has been used in the selection of electrical fixture assemblers, radio assemblers, power-sewing machine operators, watch assemblers, punch press operators, can packers, and dental students Fleishman, 1953c). It has been considered a useful test. Validity coefficients for the O'Connor test in these various programs range from −.10 to .50, with a median value of .22.

Outside of pilot selection, complex apparatus tests have not been widely used. However, in the years from 1942 to 1956, when the Air Force was using them, apparatus tests were consistently useful. The Complex Coordination Test enjoyed correlations with success or failure in primary flight training which ranged into the .40's. Other apparatus tests, for example, the Rotary Pursuit and Rudder Control Tests, were somewhat less predictive of the same criterion. Pursuit Confusion, Two Hand Coordination, and Direction Control Tests usually correlated in the .30's.

It is commonplace to observe that correlated tests do not necessarily resemble one another. The usual explanation of this phenomenon is to say that the tests involve the same "abilities." I am not sure that this "explanation" is much more than a tautology, since these "abilities" are not and, perhaps, cannot be known independently of the tests which identify them. Certainly, if two tests are positively correlated there is *something* about them which is common. It may not be suggested by an examination of the test behaviors, but ordinarily it is.

Major correlations between tests of widely disparate kinds are not a common phenomenon. In general, printed tests correlate best with other printed tests, apparatus tests with other apparatus tests, and flight behaviors with other flight behaviors. Correlations between one stage of pilot training and another are frequently above .50. Multiple correlations of early stages with later ones may range up to .90. In pilot selection the search for nonflight behaviors which have predictive validity for flight grades has been going on for many years. By now virtually everything has been tried. Nevertheless, fully 40% of the variance of stage and maneuver grades remains specific to the flight situation.

In view of this history the Navy has taken a new tack toward selective control in pilot training (Berkshire & Lyon, 1959; Shoenberger, Wherry, & Berkshire, 1963). Conventional selection is applied at the point of admission only. Nevertheless, students continue to be eliminated throughout training. This too is selection. A cadet who is dropped in pre-solo is eliminated not because he cannot fly an airplane but because in the judgment of his instructors he cannot fly one well enough to justify his continuance in the program. Implicit in this judgment is a prediction. If the cadet were continued, he would ultimately fail outright. He would not be able to land aboard a carrier or would kill himself and possibly others in the effort.

The Navy's Aviation Psychology Laboratory, under the leadership of J. R. Berkshire, formalized these predictions. At every stage of a cadet's progress, from admission through pre-solo and beyond, multiple correlations were computed with subsequent attrition from the program, using all materials which were available at that stage. At admission these materials were the selection tests. At the end of preflight they were augmented by peer ratings, academic, military, and physical fitness grades.

Multiple correlations with subsequent attrition increased from .38 at admission to .56 at the end of preflight. Since very little attrition occurs in preflight, these two coefficients are comparable; predictive efficiency at the end of preflight is twice what can be achieved before the cadets are brought into the training program; and the cost of bringing a man to Pensacola is not great in comparison with what it costs to teach him to become a naval aviator.

In a formal way, the Navy does not practice selection at the end of preflight. Instead, the probabilities of each man's completing the remainder of the program are computed. If a cadet gets into trouble in flight training, these probabilities are known and taken into consideration by the boards that decide his case.

This same procedure is followed in the remainder of the program. At the end of pre-solo, predictive validity jumps to .66. However, pre-solo is also the point of heaviest attrition from the program; if the pre-solo coefficient were corrected to make it comparable with preflight and earlier validities, it would be much higher.

The Navy has pursued this approach into the fleet. Graduates of the training program are rated by their squadron commanders as "satisfactory" or "unsatisfactory"; and multiple correlations at successive stages of training are computed against this criterion. In some instances, predictive validities as high as .70 have been obtained for fighter pilots.

The Navy's approach is far less "psychometric" than the conventional procedure. It is patterned after quality control in industry, rather than psychological theories about abilities. At the same time, however, it is entirely congruent with the findings and experience of personnel psychologists.

## Training

The psychology of learning is a laboratory discipline. It has to do with finger mazes, apparatus tests, nonsense syllables. Learning how to operate a drill press, how to fly an airplane, or how to prove mathematical theorems are not learning as the psychologist uses the word; they are training. "If an investigator's primary interest," writes Dael Wolfle (1951, p. 1267), "is in the process by which knowledge or skill is acquired, then his studies are classified under the heading of learning. But if his purpose is to investigate the teaching that goes on in school, in industry, or in a military establishment, his work is classified under the heading of training."

In this case the difference between laboratory and field is fundamental. In learning every trial is as much like every other as the experimenter can make them. In learning outside the laboratory, i.e., in training, there are no trials in the laboratory sense of the word. There are hops, courses, sessions, meetings, lessons, but no trials. The content of successive units of instruction is different. A child first learns the numbers; then he learns how to add and subtract them; then he learns the multiplication tables. Training is characterized by a constantly changing and cumulative demand.

Training is predominantly a normative area. The major questions which confront people who teach things are what to teach, how to teach it, and in what order. These are all normative questions. They have to do with schedules of training, forms of presentation, teaching techniques, and transfer of training. Nevertheless, there are individual differences in performance in all parts of any training syllabus; and one of the jobs of the differential psychologist is to describe these differences and the relations between them.

The usual approach to this problem is factor analytic. In each phase of training the trainees receive a grade or score. The correlations between each phase and every other are calculated and then factor analyzed.

Table XI consists of correlations between stages in naval air basic training. The data were collected in 1954 from the flight jackets of 300 naval aviation cadets. O. E. Rogers (1956) factored this matrix by the centroid method and rotated the factors to oblique simple structure. The analysis resulted in

TABLE XI

CORRELATIONS AMONG STAGES IN THE NAVAL AIR BASIC TRAINING SYLLABUS[a]

| *Stage* | *1* | *2* | *3* | *4* | *5* | *6* | *7* | *8* | *9* | *10* | *11* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Pre-solo | | .64 | .48 | .29 | .23 | .12 | .15 | .27 | .17 | .12 | .34 |
| 2. Precision | | | .63 | .27 | .21 | .15 | .23 | .34 | .08 | .16 | .46 |
| 3. Acrobatics | | | | .25 | .15 | .22 | .21 | .34 | .14 | .14 | .46 |
| 4. Formation | | | | | .31 | .25 | .22 | .26 | .24 | .14 | .36 |
| 5. Night primary | | | | | | .14 | .22 | .09 | .16 | .08 | .16 |
| 6. Primary combat | | | | | | | .24 | .12 | .02 | .09 | .18 |
| 7. Cross-country | | | | | | | | .08 | .02 | .09 | .16 |
| 8. Gunnery | | | | | | | | | .23 | .13 | .23 |
| 9. Field carrier landing practice | | | | | | | | | | .11 | .09 |
| 10. Carrier qualification | | | | | | | | | | | .14 |
| 11. Basic instruments | | | | | | | | | | | |

[a] After Jones (1959)

two factors. The first factor bound together formation and the primary stages, pre-solo, precision, and acrobatics. The second factor had its biggest loadings on gunnery and cross-country. Rogers called the first factor a "primary training" factor and the second a "tactics-instruments-gunnery" factor.

There are several objections to this kind of differential "description" of a training syllabus. First, the factor-analytic account is not really descriptive but theoretical. It explains the correlations by postulating the existence of two hypothetical factors with determinate relations to each other and to the various stages of training; and as an explanation, it is open to question at a theoretical level. For example, are there really only two differential factors in the whole of naval air basic training? If the task at hand is descriptive, it might be better to study the correlational results without resort to hypothetical factors; and then, if regularities are discovered, attempt a theoretical construction.

A second objection is that the factor-analytic "description" takes no account of the sequential features of training. In any training program some stages must precede others. It isn't possible, for example, to teach formation before pre-solo. Factor analysis originated in the study of psychological tests, as a taxonomic method. In other areas it isn't clear that taxonomic ideas make the most appropriate terms of description.

Both of these objections may be met by a purely descriptive account of sequential relationships among the stages of training. Consider, for example, the correlations among the first three stages, pre-solo, precision, and acrobatics. Pre-solo and acrobatics are both more strongly related to precision than they are to each other. The three correlations form a very short superdiagonal form; they form a differential sequence.

In this case either time or content could be the basis of the ordering; and as far as these three stages are concerned, it is impossible to know which is operative, because the three stages have to be taught in the order they were. The rest of the basic syllabus is more flexible. For example, night primary, primary combat, cross-country, gunnery, and basic instruments may be taught in any order; and several orders have been tried (Jones, 1959, pp. 87–88). Correlational studies of these various attempts have not found appreciable differences. The best evidence we have, therefore, suggests that correlational patterns in the naval air training syllabus are based on content.

In the case of pre-solo, precision, and acrobatics, the differential sequence appears to rest upon normative inclusions; in a general way, pre-solo behaviors are required in precision and acrobatics, and behaviors required in precision are also required in acrobatics. To confirm the hypothesis, these component behaviors would have to be enumerated and their independence established.

Consider another example. Basic instruments, formation, and gunnery also subtend correlations which have the superdiagonal form. In this case, moreover, temporal sequence cannot be the basis of the ordering because the stages were not taught in this order; in the temporal sequence gunnery was in the middle. If we eliminate all sequences which cannot be realized, because a later stage is requisite to one which precedes it in the sequence, there are four other three-stage superdiagonal forms in the matrix. In all four basic instruments and formation are the first two members.

In Fig. 3 the six sequences are presented graphically. Any string of stages (reading from left to right) which is joined together by solid lines indicates a differential sequence. These lines have no surplus meaning. They assert directly observable patterns. Stages which are serially connected by solid lines subtend correlations which are organized according to the sequential position of the stages. Each stage correlates most strongly with its immediate neighbors and most weakly with stages which are most removed from it. A broken line indicates that the immediately antecedent stage is requisite to the stage which follows it; a broken line asserts nothing about correlational pattern. This kind of graph is called a *string model.*

The most striking feature of the string model in Fig. 3 is the position it attributes to formation. Formation is the hub of the syllabus. Every stage in the basic syllabus is either antecedent or subsequent to it, as a matter of differential sequence or of normative necessity. At the same time, formation was also the most representative stage in the basic syllabus. Primary training was a preparation for formation flying; and several of the more advanced stages,

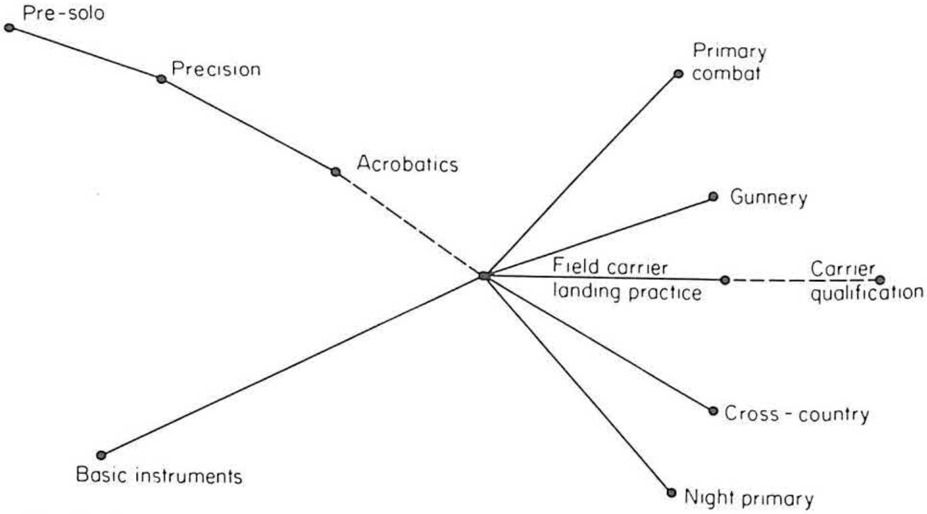


FIG. 3. A string model for the Naval air basic training syllabus; the hub stage is formation (after Jones, 1959).

for example, primary combat, gunnery, and carrier qualification were flown from formation.

Upon completion of basic, a cadet continued into jet or multiengine training. At the time these studies were done the jet syllabus contained 16 stages. The correlations among these stages contained 24 distinct superdiagonal forms, all of them normatively possible and most of them four stages long (Jones, 1959). The string model which set forth these 24 sequences had two hubs. The larger was jet formation. The second hub was an instrument stage, direction finding; and it was differentially antecedent to formation. This second hub reflected the introduction into the jet syllabus of more instrument stages, such as ground control intercept, radio range, and ground control approach.

In the multiengine syllabus there was no formation stage; aside from familiarization to the larger planes, the syllabus consisted exclusively of instrument stages. The string model for this syllabus had a single hub in basic instruments.

In basic, jet, and multiengine training the normative and differential organization of the syllabuses were essentially parallel. Differential sequences generally reflected a progression from basic to more advanced skills; and stages which were central to the content of the syllabus were hubs in its differential patterning.

Sequential descriptions may also be made of correlations among different aspects of a single stage of training. Among the many studies made by the

Army Air Force during World War II was a correlational analysis of primary pilot training (Ben-Avi, 1947). It was based on the grades of 500 cadets in each of 15 different maneuvers. These 15 maneuvers were not taught in a linear order as were the stages of naval air training; they were taught intermittently throughout the stage. Of course, some maneuvers, like straight and level flying and take-offs, were taken up before others, like 8's along a road or forced landings. But instruction in the more elementary maneuvers usually continued after the more complicated ones had been introduced.

In Fig. 4 a string model for the 15 maneuvers is presented. It contains five differential sequences. They begin with straight and level flying, taxiing, or coordination exercises. Three sequences end in forced landings; the other two end in mid-air turning maneuvers.

These sequences are plainly parallel to other progressions which even nonaviators can appreciate. For example, with the possible exception of the last maneuver, taxiing, take-offs, landings, and forced landings are an aeronautic succession. A pilot must taxi before he can take off, and he must take off before he can land. Other sequences, for example, straight and level flying, turns, 8's across a road, and rectangular courses clearly reflect the building of more complicated skills upon simpler ones.

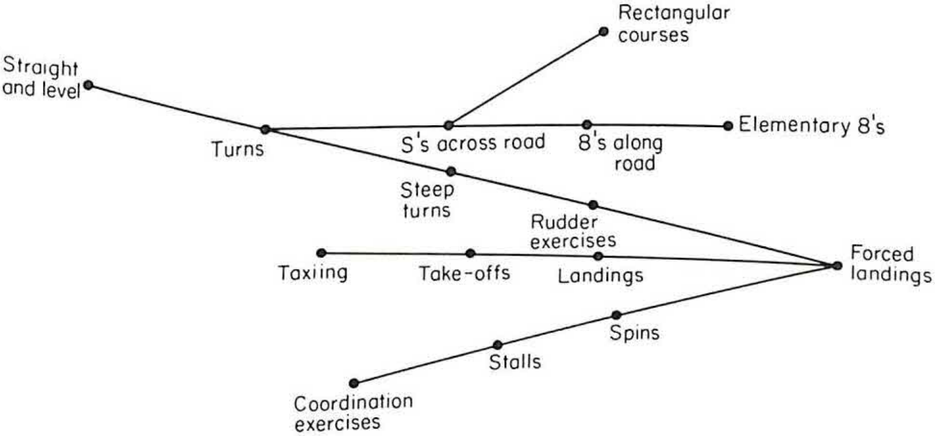


FIG. 4. A string model for maneuvers in primary pilot training.

## Tight Little Islands

As practice proceeds on a psychomotor task the proportion of variance which is specific to the task, i.e., which cannot at present be related to external variables, increases. In the Complex Coordination Test, for example, the proportion of specific variance increases from 6 percent at the beginning to 38 percent at the end of practice (Fleishman & Hempel, 1953). In psychomotor tests generally, most of the variance at the end of practice is specific to the task. This finding has tended to focus attention for the first time in many years on specific variance.

In one analysis of psychomotor test relationships (Fleishman, 1954) the proportion of variance which was specific to the tests averaged 49 percent. This result is typical. In field and criterion behaviors the proportion of specific variance is much larger. Usually, at least 75 percent of the variance is specific to the tasks, frequently as much as 90 percent.

In his presidential address to the APA, Cronbach characterized experimental psychology as a "tight little island" and differential psychology as a "Holy Roman Empire." In the case of experimental psychology the plural might have been more descriptive. In any experiment there is a great deal of unexplained variance within experimental conditions. Known experimental relationships exist as many islands in a sea of within-group error. And the same is true of differential psychology. Its findings are so many islands surrounded on all sides by expanses of specific variance.

## REFERENCES

Adams, J. A. The prediction of performance at advanced stages of training on a complex psychomotor task. *USAF Hum. Resour. Res. Center res. Bull.*, 1953, No. 53–49.

Adams, J. A. The relationship between certain measures of ability and the acquisition of a psychomotor criterion response. *J. gen. Psychol.*, 1957, **56**, 121–134.

Ben-Avi, A. H. Studies of subjective measures of flying proficiency. In N. E. Miller (Ed.), *Psychological research on pilot training*. Washington, D.C.: U. S. Govt. Printing Office, 1947. Pp. 83–109.

Berkshire, J. R., & Lyon, V. W. Human quality control in naval air training. *Amer. Psychologist*, 1959, **14**, 153–155.

Bilodeau, E. A. Transfer of training between tasks differing in degree of physical restriction of imprecise responses. *USAF Hum. Resour. Res. Center res. Bull.*, 1952, No. 52–40.

Bilodeau, E. A. Variations in knowledge of component performance and its effects upon part-part and part-whole relations. *J. exp. Psychol.*, 1955, **50**, 215–224.

Bilodeau, E. A. Patterns of internal consistency in multipart skilled performances. *Amer. J. Psychol.*, 1957, **70**, 550–559.

Bilodeau, E. A., Jones, M. B., & Levy, C. M. Long-term memory as a function of retention time and repeated recalling. *J. exp. Psychol.*, 1964, **67**, 303–309.

Borgatta, E. F. Sidesteps toward a nonspecial theory. *Psychol. Rev.*, 1954, **61**, 343–352.

Brody, D. Twin resemblances in mechanical ability, with reference to the effects of practice on performance. *Child Develpm.*, 1937, **8**, 207–216.

Cronbach, L. J. The two disciplines of scientific psychology. *Amer. Psychologist*, 1957, **12**, 671–684.

Edgerton, H. A., & Valentine, W. L. A factor analysis of learning data. *Psychol. Bull.*, 1935, **32**, 719.

Fleishman, E. A. A factor analysis of intra-task performance on two psychomotor tests. *Psychometrika*, 1953, **18**, 45–55. (a)

Fleishman, E. A. Testing for psychomotor abilities by means of apparatus tests. *Psychol. Bull.*, 1953, **50**, 241–262. (b)

Fleishman, E. A. A modified administration procedure for the O'Connor Finger Dexterity Test. *J. appl. Psychol.*, 1953, **37**, 191–194. (c)

Fleishman, E. A. A factorial study of psychomotor abilities. *USAF Personnel Train. Res. Center res. Bull.*, 1954, No. 54–15.

Fleishman, E. A. Factor structure in relation to task difficulty in psychomotor performance. *Educ. psychol. Measur.*, 1957, **17**, 522–532.

Fleishman, E. A. Abilities at different stages of practice in rotary pursuit performance. *J. exp. Psychol.*, 1960, **60**, 162–171.

Fleishman, E. A., & Hempel, W. E., Jr. Changes in factor structure of a complex psychomotor test as a function of practice. *USAF Hum. Resour. Res. Cent. res. Bull.*, 1953, No. 53–68.

Fleishman, E. A., & Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. *J. exp. Psychol.*, 1955, **49**, 301–312.

Fleishman, E. A., & Parker, J. F., Jr. Prediction of advanced levels of proficiency in a complex tracking task. *USAF WADC tech. Rep.*, 1959, No. 59–255.

Fleishman, E. A., & Parker, J. F., Jr. Factors in the retention and relearning of perceptual-motor skill. *J. exp. Psychol.*, 1962, **63**, 215–226.

Freedman, D. G., & Keller, B. Inheritance of behavior in infants. *Science*, 1963, **140**, 196–198.

Galton, F. *Hereditary genius.* New York: Appleton, 1887.

Greene, E. B. An analysis of random and systematic changes with practice. *Psychometrika*, 1943, **8**, 37–53.

Guilford, J. P. The structure of intellect. *Psychol. Bull.*, 1956, **53**, 267–293.

Guttman, L. A new approach to factor analysis: the radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences.* Glencoe, Ill.: Free Press, 1954. Pp. 258–348.

Houston, R. C. An evaluation of the predictive properties of measures of variability of performance on three psychomotor tasks. Unpublished doctoral dissertation, Univer. of Maryland, 1950.

Jones, M. B. Simplex theory. *USN Sch. aviat. Med. Monogr.* 1959, No. 3.

Jones, M. B. Molar correlational analysis. *USN Sch. aviat. Med. Monogr.* 1960, No. 4.

Jones, M. B. Practice as a process of simplification. *Psychol. Rev.*, 1962, **69**, 274–294.

McNemar, Q. Twin resemblances in motor skills, and the effect of practice thereon. *J. genet. Psychol.*, 1933, **42**, 70–97.

Melton, A. W., (Ed.), *Apparatus tests.* Washington, D.C.: U. S. Govt. Printing Office, 1947.

Perl, R. E. An application of Thurstone's method of factor analysis to practice series. *J. gen. Psychol.*, 1934, **11**, 209–212.

Reynolds, B. Correlation between two psychomotor tasks as a function of distribution of practice on the first. *J. exp. Psychol.*, 1952, **43**, 341–348. (a)

Reynolds, B. The effect of learning on the predictability of psychomotor performance. *J. exp. Psychol.*, 1952, **44**, 189–198. (b)

Rogers, O. E. Analyses of basic training stage grades for multi-engine and single-engine aviators. *USN Sch. aviat. Med. res. Rep.*, 1956, No. 3.

Schoenberger, R. W., Wherry, R. J., Jr., & Berkshire, J. R. Predicting success in aviation training. *USN Sch. aviat. Med. res. Rep.*, 1963, No. 7.

Spearman, C. General intelligence, objectively determined and measured. *Amer. J. Psychol.*, 1904, **15**, 201–293.

Thomas, A., Chess, S., Birch, H. G., Hertzig, M. E., & Korn, S. *Behavioral individuality in early childhood.* New York: New York Univer. Press, 1964.

Viteles, M. S. The influence of training on motor test performance. *J. exp. Psychol.*, 1933, **16**, 556–564.

Wolfle, D. Training. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. p. 1267.

Woodrow, H. The effect of practice on groups of different initial ability. *J. educ. Psychol.*, 1938, **29**, 268–278.

Woodrow, H. Factors in improvement with practice. *J. Psychol.*, 1939, **7**, 55–70. (a)

Woodrow, H. The application of factor analysis to problems of practice. *J. gen. Psychol.*, 1939, **21**, 457–460. (b)

# Human Abilities and the Acquisition of Skill

## Comments on Professor Jones' Paper

EDWIN A. FLEISHMAN
*American Institutes for Research*
*Washington, D.C.*

It is encouraging to see the topic of individual differences included in a conference on skill acquisition. It was not too long ago when learning psychologists viewed individual differences as troublesome error variance, obscuring rather than clarifying the principles of skill learning. Yet our own program indicates that we can learn much about human learning if we study the relationships between the ways in which people differ and their learning of complex skills. Jones has presented a very stimulating, albeit off-beat, paper on individual differences, and I am glad to have the opportunity to discuss it. I have also been asked to use this opportunity to put on the record, and into this book, some of my views and data on motor abilities and performance.

My objectives, which I cannot possibly accomplish, are to discuss, then, the issues raised by Jones, to present additional substantive data on individual differences and motor abilities, and to try to keep the "state of art" needs in mind. One approach is to try to take up, in turn, the issues as Jones sees them. In view of my multiple objectives here, I prefer to emphasize whatever substantive contributions I can, dealing with the issues raised by Jones wherever I can, as they arise.

### The Ability-Skill Distinction

First, I would like to discuss the concepts of ability and skill. Jones seems not to like the term *abilities*, but abilities would seem to have the same status as many other constructs in psychology, and I have no objection at all to using the term. In my own work I have found it convenient to distinguish between the terms *ability* and *skill*.

As we use the term, *ability* refers to a more general trait of the individual which has been inferred from certain response consistencies (e.g., correlations) on certain kinds of tasks. These are fairly enduring traits, which in the adult, are more difficult to change. Many of these abilities are, of course, themselves a product of learning, and develop at different rates, mainly during childhood and adolescence. Some abilities (e.g., color vision) depend more on genetic than learning factors, but most abilities depend on both to some degree. In any case, at a given stage of life, they represent traits or organismic factors which the individual brings with him when he begins to learn a new task. These abilities are related to performances in a variety of human tasks. For example, the fact that spatial visualization has been found related to performance on such diverse tasks as aerial navigation, blue-print reading, and dentistry makes this ability somehow more basic.

The term *skill* refers to the level of proficiency on a specific task or limited group of tasks. As we use the term *skill*, it is task oriented. When we talk about proficiency in flying an airplane, in operating a turret lathe, or in playing basketball, we are talking about a specific skill. Thus, when we speak of acquiring the skill of operating a turret lathe, we mean that this person has acquired the sequence of responses required by this specific task. The assumption is that the skills involved in complex activities can be described in terms of the more basic abilities. For example, the level of performance a man can attain on a turret lathe may depend on his basic abilities of manual dexterity and motor coordination. However, these same basic abilities may be important to proficiency in other skills as well. Thus, manual dexterity is needed in assembling electrical components and motor coordination is needed to fly an airplane. There are, of course, certain characteristics of skill on which we could elaborate. Thus most skills involve some (a) spatial-temporal patterning, (b) interaction of responses with input and feedback processes, and (c) learning.

## Abilities and Skill Learning

Implicit in the previous analysis is the important relation between abilities and learning. Thus, individuals with high manual dexterity may more readily learn the specific skill of lathe operation. The mechanism of transfer of training probably operates here. Some abilities may transfer to the learning of a greater variety of specific tasks than others. In our culture, *verbal* abilities are more important in a greater variety of tasks than are some other types of abilities. The individual who has a great many highly developed basic abilities can become proficient at a great variety of specific tasks. The concept of intelligence really refers to a combination of certain basic abilities which contribute to achievement in a wide range of different activities.

Elsewhere (Fleishman, 1964; Gagné & Fleishman, 1959) we have elaborated our analysis of the development of basic abilities. This included a discussion of their physiological bases, the role of learning, environmental, and cultural factors, and evidence on the rate of ability development during the life span. Particularly critical is the notion that basic abilities are relatively enduring traits of the individual. Unless he is subjected to marked environmental changes, a man's basic abilities are not likely to change much once he reaches adulthood. Thus, such abilities as numerical, verbal, and spatial abilities are fairly stable attributes of behavior over lengthy periods of time. This probably results from the fact that they have been learned, relearned, and practiced so many times during the individual's lifetime. The mechanism here is probably what we typically call overlearning. Thus, individuals may show no impairment of performance even after lengthy periods of time without practice. Probably, man's abilities are stable because certain kinds of human performances have been practiced in varying degrees throughout an individual's lifetime and have become "overlearned."

Elsewhere (Gagné & Fleishman, 1959) I have presented cross-sectional ability comparisons for ages 17–60. Typically, these curves indicate that no improvement and very little fluctuation in the levels of abilities occur during much of the adult life span. Some decline with age often occurs, but this is true only of some abilities, not all. And the age at which decline begins and the rate of decline vary with the ability. Thus, verbal ability shows virtually no decline, numerical ability declines after age 50, but spatial ability and memory span decline after age 40, the spatial curve dropping more steeply.

Human adults, of course, show marked learning over time in practically any type of specific *skill*. However, according to our conceptualization, the rate of learning and the final level achieved by particular individuals in certain skills are both limited by the basic abilities of these individuals. The fact that these basic abilities are themselves fairly stable allows us to make useful predictions about subsequent performance in specific tasks. For example, knowledge about a person's numerical ability helps us predict his probable success later on in engineering training. Knowledge about the relevant physical fitness components should help us predict performance in complex athletic skills. We will illustrate this later, but let us say now that it is very clear that knowledge about basic abilities helps us predict subsequent more complex performances. We also see that some abilities may be more important later in skills learning, while other abilities are critical earlier in learning. The idea that basic abilities place limits on later skill proficiency emphasizes the need to develop these abilities in pre-adult life.

With this conceptualization in mind, we can say that in much of my own previous work one objective has been to describe certain skills in terms of these more general ability components. Jones feels the term "abilities" recalls to him a tradition which goes back to instinct and faculty psychology. I do not see his point about instinct, but there is certainly a similarity to the older faculty notions. The difference is that the present "abilities" are empirically derived and not based on arm-chair speculation. I would say that the identification of human abilities is more akin to the old functional psychology. I want to return to Jones' views on abilities and their identification a little later on.

First, I'd like to describe some of my own work as it relates to individual differences in motor abilities and performance. This may place Jones' views in a context in which some of our agreements and differences can be underscored. And it illustrates a research strategy, not brought out in Jones' paper, which has guided our research in this area.

## Taxonomy of Perceptual-Motor Abilities

The first area of our work relates to the taxonomy problem discussed by Jones. We both see the need to develop a systematic classification of human tasks and I was very happy to see Jones bring this out as a critical problem. I have felt that the extent to which we can generalize research results is, in some measure, dependent on the development of an adequate taxonomy. This is also needed to assist in the standardization of experimental tasks. At the start of our program in 1951, we did, indeed, make the development of a taxonomy of perceptual-motor abilities our first order of business.

The impetus for this basic research program was a very applied problem. While with the Air Force Personnel and Training Research Center, one of our missions was to build better psychomotor tests for the prediction of pilot success. The wartime Air Force program had been highly successful in developing such tests (Melton, 1947). For example, the Complex Coordination Test, pictured in Fig. 1c, had consistent validity for pilots. This seems not surprising, since the test has high face validity for the pilot's job. The pilot does manipulate stick and rudder controls. But there were many tests which had substantial validity, but did not at all "resemble" the pilot's job. Cases in point are the Two Hand Coordination and Rotary Pursuit Tests (Figs. 1a and 1b). So it seemed to me that the first step was to discover the sources of validity in these tests. What ability factors were there in common to psychomotor tests which were common to the pilot criterion? If this were known, we could attempt to maximize this validity in future tests. In subsequent years we have conducted a whole series of interlocking experimental–factor–analytic studies, attempting to isolate and identify the common variance in a wide range of psychomotor performances. Essentially this is laboratory research, in which tasks are specifically designed or selected to test a certain hypothesis about the organization of abilities in a certain range of tasks. Subsequent studies tend to introduce task variations aimed at sharpening or limiting our ability factor definitions. The purpose is to define the fewest independent ability categories which might be most useful and meaningful in describing performance in the widest variety of tasks.

Our first study (Fleishman, 1953b) reviewed previous factor analyses of motor skills and described the methodological issues. Our second study (Fleishman, 1954) was a large-scale attempt to put into a single study representative measures of factors previously identified. The purpose of this study, which involved more than 40 carefully designed psychomotor tests, was to (a) better define previously identified factors, (b) test the usefulness of certain printed measures of psychomotor abilities, and (c) to attempt to account for variance in tests used to select pilots in the Air Force program. In short, this study confirmed most of the hypothesized factors and resulted in better definitions of these factors. But there were some exceptions and new discoveries. We also showed that printed tests are not too successful in reproducing variance in apparatus tests, and we failed to identify the variance in the complex Air Force tests. For the latter result, we had to look elsewhere.

Jones' paper, in the section on taxonomy, uses these 1953–1954 papers as the basis for his summary of psychomotor factors. Actually, these two papers represented only a starting point; the main substantive work followed, building on these earlier papers. Hence, I would not like to see Jones' list immortalized in this volume as the definitive list of psychomotor factors. As we shall see, some of the most general factors were identified only later.

Our later studies have included analyses of fine manipulative performances (e.g., finger and manual dexterity, etc.) (Fleishman & Hempel, 1954a; Fleishman & Ellison, 1962), and gross physical proficiency (e.g., push-ups, chins, etc.) (Hempel & Fleishman, 1955; Nicks & Fleishman, 1962; Fleishman, Shoup, & Kremer, 1961; Fleishman, Thomas, & Munroe, 1961; Fleishman 1963; Fleishman, 1964). One study focused on positioning movements (reaching, moving controls to specified positions, etc.) and "static reactions" (e.g., hand steadiness) (Fleishman, 1958a). The former concerns movements in which the terminal accuracy of the response is critical, and the latter involves primarily maintenance of limb positions (see Brown & Jenkins, 1947). The most important series of studies (e.g., Fleishman, 1956; Fleishman, 1958b; Parker & Fleishman, 1960) concerned "movement reactions," where the performance involves coordinated responses, or smooth responses, or precisely controlled movements, or continuously adjustive reactions.

Thus far, we have investigated more than 200 different tasks administered to thousands of subjects in a series of interlocking studies. From the patterns of correlations obtained, we have been able to account for performance on this wide range of tasks in terms of a relatively small number of abilities. In subsequent studies, our definitions of these abilities and their distinctions from one another are becoming more clearly delineated. Furthermore, it is now possible to specify the tests which should provide the best measure of each of the abilities identified. Figure 1 illustrates some representative measures used in one of these studies.

Let us list the more important ability factors identified. For details of their definition and the devices which best measure them the reader is referred elsewhere (Fleishman, 1960b, 1964).

*Control precision:* This factor is common to tasks which require fine, highly controlled, but not overcontrolled, muscular adjustments, primarily where larger muscle groups are involved (Fleishman, 1958b; Fleishman & Hempel, 1956; Parker & Fleishman, 1960). This ability extends to arm-hand as well as to leg movements. It is most critical where such adjustments must be rapid, but precise.

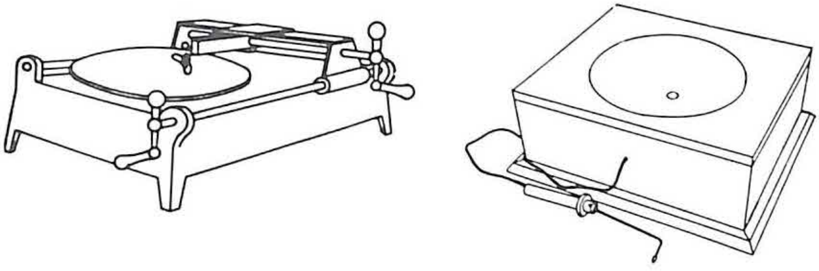
*Multilimb coordination:* This is the ability to coordinate the movements of a number of limbs simultaneously, and is best measured by devices involving multiple controls (Fleishman, 1958b; Fleishman & Hempel, 1956; Parker & Fleishman, 1960). The factor has been found general to tasks requiring coordination of the two feet (e.g., the Rudder Control Test), two hands (e.g., the Two Hand Pursuit and Two Hand Coordination Tests), and hands and feet (e.g., the Plane Control and Complex Coordination Tests).

*Response orientation:* This ability factor has been found general to visual discrimination reaction psychomotor tasks involving rapid directional discrimination and orientation of movement patterns (Fleishman, 1957a,b, 1958; Fleishman & Hempel, 1956; Parker & Fleishman, 1960). It appears to involve the ability to *select* the correct movement in relation to the correct stimulus, especially under highly speeded conditions.

*Reaction time:* This represents simply the speed with which the individual is able to respond to a stimulus when it appears (Fleishman, 1954, 1958b; Fleishman & Hempel, 1955; Parker & Fleishman, 1960). There are consistent indications that individual differences in this ability are independent of whether the stimulus is auditory or visual and are also independent of the type of response which is required. However, once the stimulus situation or the response situation is complicated to involve alternate choices, reaction time is not the primary factor that is measured.

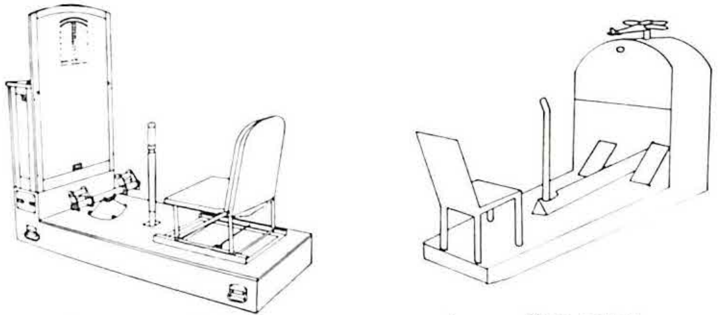
*Speed of arm movement:* This represents simply the speed with which an individual can make a gross, discreet arm movement where accuracy is not the requirement (Fleishman, 1958b; Fleishman & Hempel, 1954b, 1955; Parker & Fleishman, 1960). There is ample evidence that this factor is independent of the reaction time factor.

*Rate control:* This ability involves the making of continuous anticipatory motor adjustments relative to changes in speed and direction of a continuously moving target or object (Fleishman, 1958; Fleishman & Hempel, 1955, 1956). This factor is general to tasks involving compensatory as well as following pursuit, and extends to tasks involving responses to changes in rate. Our research has shown that adequate measure-



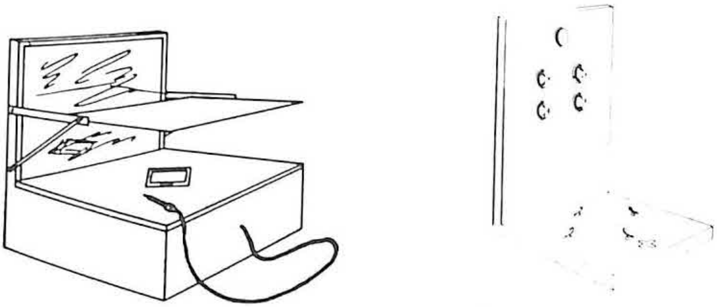
a Two hand coordination

b Rotary pursuit



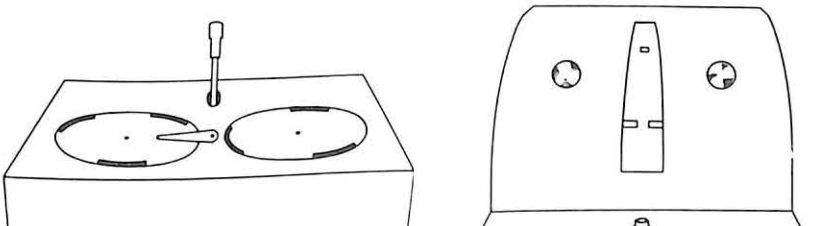
c Complex coordination

d Plane control



e Pursuit confusion

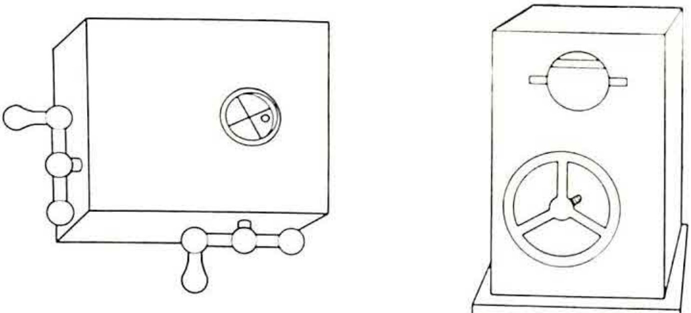
f Discrimination reaction time



g Motor judgment

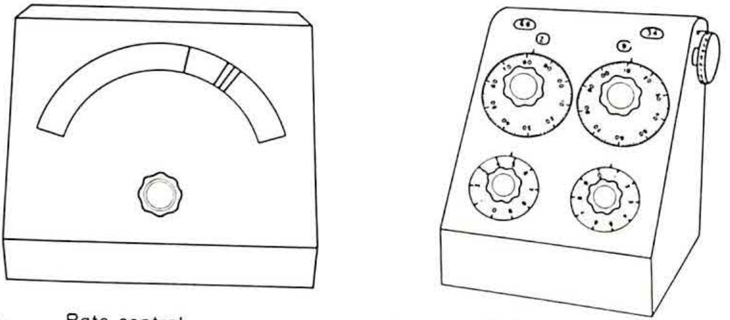
h Visual coincidence

FIG. 1. Some testing devices used in research on the classification of perceptual-motor abilities. From Fleishman (1958b).



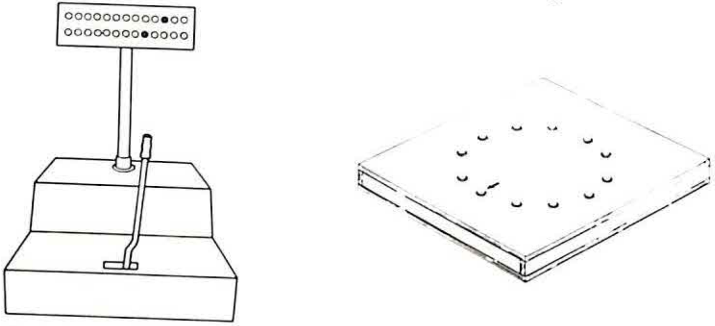
i. Two hand pursuit

j Single dimensional pursuit



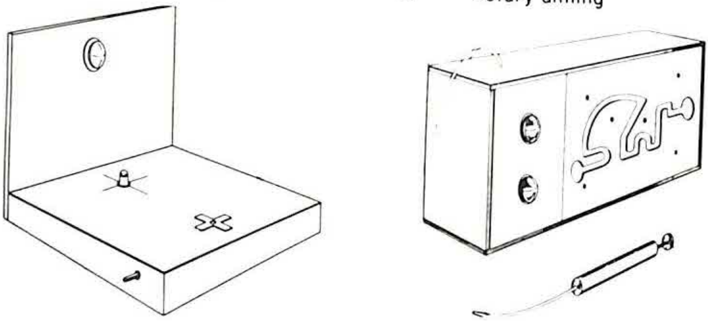
k. Rate control

l. Dial setting



m. Controls adjustment

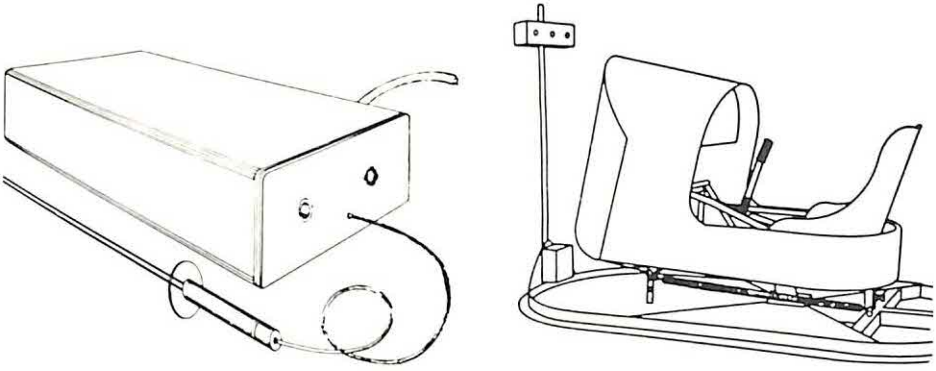
n. Rotary aiming



o. Reaction time

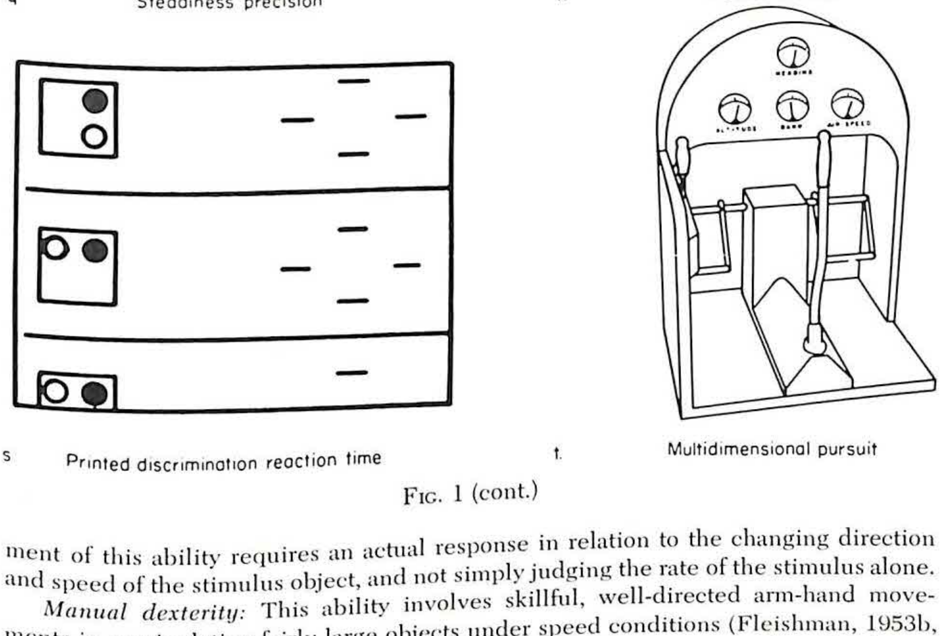
p. Track tracing

Fig. 1 (cont.)



q Steadiness precision

r. Rudder control



s Printed discrimination reaction time

t. Multidimensional pursuit

FIG. 1 (cont.)

ment of this ability requires an actual response in relation to the changing direction and speed of the stimulus object, and not simply judging the rate of the stimulus alone.

*Manual dexterity:* This ability involves skillful, well-directed arm-hand movements in manipulating fairly large objects under speed conditions (Fleishman, 1953b, 1954; Fleishman & Hempel, 1954b; Parker & Fleishman, 1960; Hempel & Fleishman, 1955; Fleishman & Ellison, 1962).

*Finger dexterity:* This is the ability to make skill-controlled manipulations of tiny objects involving, primarily, the fingers (Fleishman, 1953b; 1954; Fleishman & Hempel, 1954b; Parker & Fleishman, 1960; Hempel & Fleishman, 1955; Fleishman & Ellison, Ellison, 1962).

*Arm-hand steadiness:* This is the ability to make precise arm-hand positioning movements where strength and speed are minimized; the critical feature, as the name implies, is the steadiness with which such movements can be made (Fleishman, 1953b, 1954, 1958a,b; Fleishman & Hempel, 1955; Hempel & Fleishman, 1955; Parker & Fleishman, 1960).

*Wrist, finger speed:* This ability has been called "tapping" in many previous studies through the years (e.g., Greene, 1943; Fleishman, 1953). It has been used in a variety of different studies, primarily because these are in the form of printed tests which are quick and easy to administer. However, our research shows that this factor is highly restricted in scope and does not extend to many tasks in which apparatus is used (Fleishman, 1954; Fleishman & Hempel, 1954b). It has been found that the factor is best measured by printed tests requiring rapid tapping of the pencil in relatively large areas.

*Aiming:* This ability appears to be measured by printed tests which provide the subject with very small circles to be dotted in where there are a large number of circles when the test is highly speeded (Fleishman, 1953, 1954; Hempel & Fleishman, 1955). The subject typically goes from circle to circle placing one dot in each circle as rapidly as possible. This factor has not been found to extend to apparatus tests and, hence, the naming of this factor as "aiming" or as other investigators have called it "eye-hand coordination" seems much too broad (Fleishman & Ellison, 1962).

The usefulness of these ability factors in describing skill in complex, "real-world" tasks has been described elsewhere (e.g., Fleishman & Ornstein, 1960; Herbert, 1963; Locke, Zavala, & Fleishman, 1965). Work under way is attempting to define laboratory test variables in terms of these categories (e.g., in our current studies on the effects of drugs on human performance).

Before closing our discussion of the classification of motor abilities we should refer to the area of motor performance often called *physical proficiency*. Our experimental–factor-analytical work indicates that the following factors account for performance in more than 60 different physical fitness tests. [For details, see Fleishman (1964); this reference also indicates the tests most diagnostic of each factor.]

*Extent flexibility:* Ability to flex or stretch the trunk and back muscles *as far as possible* in a forward, lateral, or backward direction.

*Dynamic flexibility:* The ability to make repeated, *rapid* flexing movements in which the resiliency of the muscles in *recovery* from strain or distortion is critical.

*Static strength:* The maximum *force* which a subject can exert, for a brief period, where the force is exerted continuously up to this maximum. In contrast to other strength factors, this is the force which can be exerted against external objects (e.g., lifting heavy weights, pulling against a dynamometer), rather than in supporting or propelling the body's own weight.

*Dynamic strength:* The ability to exert muscular force repeatedly or continuously over time. It represents muscular endurance and emphasizes the resistance of the muscles to fatigue. The common emphasis of tests measuring this factor is on the power of the muscles to propel, support, or move the body repeatedly or to support it for prolonged periods.

*Trunk strength:* This is a second, more limited, dynamic strength factor specific to the trunk muscles, particularly the abdominal muscles.

*Gross body coordination:* Ability to coordinate the simultaneous actions of different parts of the body while making gross body movements.

*Gross body equilibrium:* The ability of an individual to maintain his equilibrium, despite forces pulling him off balance, where he has to depend mainly on nonvisual (e.g., vestibular and kinesthetic) cues. Although also measured by balance tests where the eyes are kept open, it is best measured by balance tests conducted with the eyes closed.

*Stamina:* The capacity to continue maximum effort, requiring prolonged exertion over time. This factor has the alternate name of "cardiovascular endurance."

Let us return to some of Jones' comments on the taxonomy question and to the problem of quantifying human abilities in general. In the early section of his paper he refers to the cleavage between experimental and differential psychology as virtually absolute. I think this has been all too true, as decried by Cronbach (1957), Ferguson (1954, 1956), and others. That this is certainly not universally true is evidenced in my own work, and in the work of several others of the participants of this conference. I like to think that the work I have described (and will describe in the next section) does bridge experimental and differential methods. But there are some unkind cuts in Jones' paper. He calls differential psychology "passive, speculative, and leisurely. Its data are given; they are found, not produced." They are "arm-chair." In contrast, experimentalists are "active" and "hard-boiled" with "worlds waiting to be conquered." I suppose I have not been aware of this distinction before! He also points out that the task of the differential psychologist is to "measure" abilities. "If he succeeds, then the whole domain of psychomotor learning has been properly described and accounted for." I don't know of any psychologist besides Jones who has made such statements.

The problem of proliferation of factors is another problem that concerns Jones, since it is very clear that more and more factors continue to be discovered. But is this unique to the psychological world? I think it is a common scientific phenomenon. And the fact that mental life is complex should not startle any of us. A more legitimate question here concerns the proportion of common variance we are able to account for in terms of a relatively small number of abilities identified. It does appear, in the perceptual-motor domain, as I have tried to describe, that a relatively small number are very useful in organizing quite meaningfully a wide variety of performances. This does not mean that there are no more factors to discover. And it does not mean these factors identified account for a large proportion of the variance in every psychomotor task. Far from it, as we shall see later. There is much specificity, but the pursuit of the common variance is one of the primary tasks of the psychologist.

I do not really understand the full implications of Jones' assertion that correlations among tests reflect the organization of the tests and not of the individuals who take them. This seems to me to miss the point. When the experimental psychologist studies the effect of different kinds of stimuli on reaction time, his results also tell us little about the people who take them. What we learn through the arrangements of the experimental conditions and through the response measures we obtain from people, is something about a set of functional relationships and we infer something about the human processes involved. This is no less true in the study of human abilities.

I was glad to see Jones discuss E. A. Bilodeau's important work on skill composition, through his analyses of intratask correlations of the multidimensional pursuit apparatus. There are certainly alternative ways of looking at skill composition, and some are more useful for some purposes than others. Jones discussed this study under the heading of task classification. I am not sure this particular study provides a basis for task classification, although it does contribute to our understanding of some of the rules of part-whole task relationships. The problem of generalizing or predicting total task performance from prior practice on task components is a bit more complicated as I have pointed out elsewhere (Fleishman, 1965). The ability-skill paradigm I described earlier has been helpful in this regard.

## Skill Acquisition

Jones has contributed a major portion of his paper to individual differences in *skill acquisition*. This, as may be recalled, represents the second main line of our research which I want to review. Our work has focused on the problem of predicting performance at different stages of learning, and especially at eventual high levels of proficiency in complex tasks.

Jones is enamored with the oft found result that correlation matrices among learning trials generally form a pattern called the *superdiagonal form*. In such matrices each trial tends to correlate highest with the trials that immediately precede it and correlations tend to decrease the greater the separation between trials. In addition, Jones has usefully shown that sequences of four trials may follow Spearman's law of single tetrad differences. In an eight-trial matrix there are 70 such equations and Jones has shown these to approximate zero within some magnitude of error. Jones stresses the universality of this finding, and there is no question that most learning trial matrices in the literature have this superdiagonal form, although many are far from satisfying the tetrad difference criterion.

The critical point is that there are alternative ways of explaining these findings and Jones prefers a simplicial hypothesis. This means that trial 1 involves all the common factors found in subsequent trials and that as practice continues, common factors drop out so that the last trial involves only one of the factors found in trial 1. However, these factors are composites to begin with, so we really can't identify what they are. But the theory fits the data so beautifully that we do not need to look further, in Jones view. We can simply think of practice as a process of elimination.

As Jones stated in his paper, he and I are in general agreement with respect to the increase in specificity found as practice on motor-skill tasks continues. But I feel this is an *incomplete* description of what is happening in the process of skill learning.

Stopping with a simplicial analysis doesn't seem to lead us far enough along in the development of new concepts which will organize existing data more meaningfully, stimulate new experiments, or lead to improved predictions or control in new learning situations.

Let me illustrate by describing the strategy of our research in this area, which, I must openly confess, has used (but is not confined to) the factor-analytic approach. As our confidence in our concepts and variables has increased, we have relied more on multiple regression and a variety of experimental approaches. First, let me acknowledge my debt to the earlier study by Adams (1953) and before that, one by Bradley Reynolds (1952). Jones' paper has usefully pointed out still earlier work by Perl, by Woodrow, and by Edgerton and Valentine. Many of these earlier studies used fairly brief practice periods, paper and pencil tests, and relatively unsophisticated reference measures. Except for Woodrow they were not programmatic efforts. We have been fortunate in being able to pursue the general problem in a fairly persistent fashion over the past dozen years.

Essentially, our studies have attempted to relate the ability factors previously described to performance at different stages of learning more complex skills. Special interest was in predicting eventual high-proficiency levels in such skills. One of our early studies was confined to the analyses of intertrial correlations of two similar tasks practiced in different orders (Fleishman, 1953a), but subsequent studies have always included "reference measures" external to the practice task. In a typical study 200–300 subjects received a battery of reference tests known to sample certain abilities and then received practice on a criterion practice task. Through the use of factor-analytic techniques of the correlation patterns obtained, we could examine the loadings of successive trial scores on the criterion task on the factors defined by the reference tasks (Fleishman, 1960a; Fleishman & Hempel, 1954a, 1955).

In general, these studies, with a great variety of practice tasks, show that (a) the particular combination of abilities contributing to performance changes as practice continues; (b) these changes are progressive and systematic and eventually become stabilized; (c) the contribution of "nonmotor" abilities (e.g., verbal, spatial), which may play a role early in learning, decreases systematically with practice, relative to "motor abilities"; and (d) there is also an increase in a factor specific to the task itself. Figure 2 illustrates a typical result obtained using a visual discrimination reaction task.

Although in earlier papers Bechtoldt (1962) and Jones (1962) have been critical of our factor analysis of matrices containing both reference and learning trial variables, they have only to read our published work to see that a variety of different approaches have been used, all confirming these basic findings. Additional designs have used separate cross-sectional analyses of

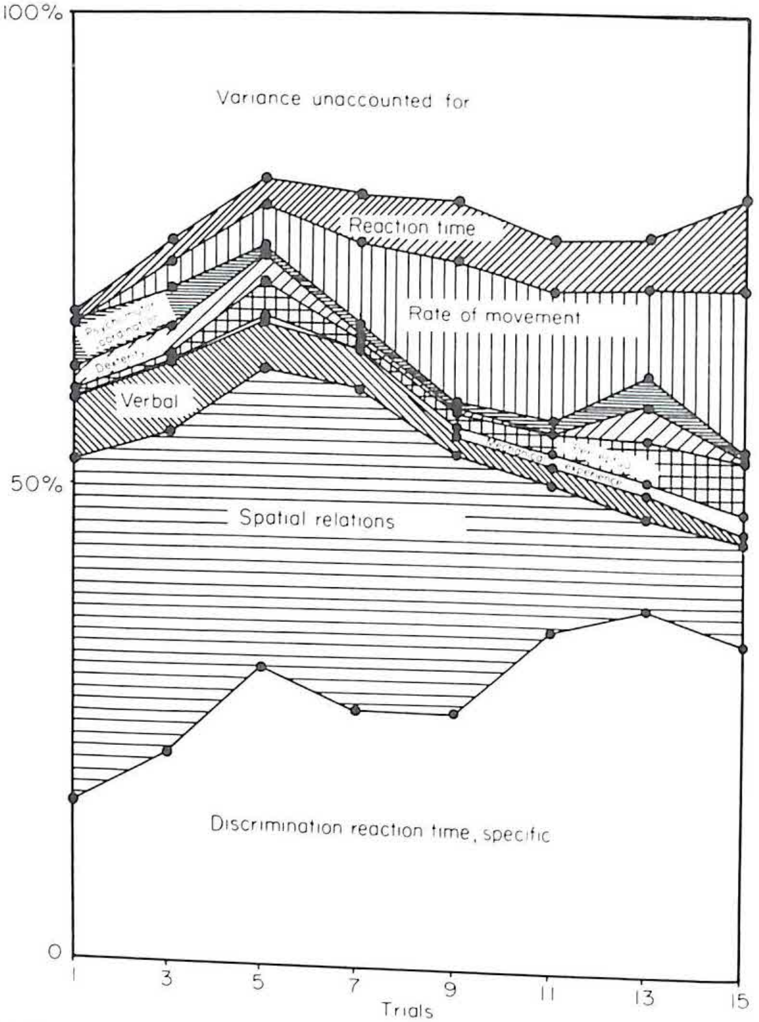


FIG. 2. Percentage of variance represented by each factor at different stages of practice on the Discrimination Reaction Time Task (Percentage of variance is represented by the *area* shaded in for each factor.) From Fleishman & Hempel (1956).

skilled and unskilled psychomotor performances (Fleishman, 1957), regression techniques in predicting practice trial loadings or reference factors (Parker & Fleishman, 1960; Fleishman & Fruchter, 1960), analyses of component-total task relations at different stages of practice (Fleishman & Fruchter, 1965), and recent work has used straight experimental procedures, as we gained confidence in our individual difference variables (Fleishman & Rich, 1963). It is possible, of course, to represent some of the results in Fig. 2 in more traditional terms. This is shown in Fig. 3.

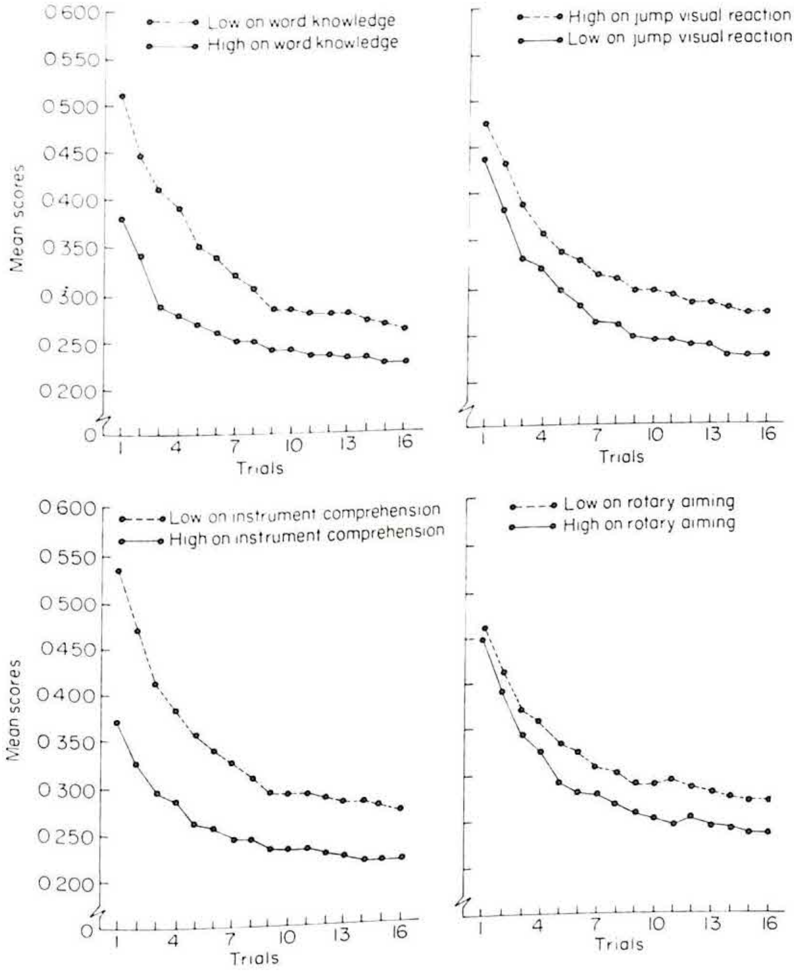


FIG. 3. Comparison of Discrimination Reaction Time acquisition curves for groups stratified on different ability test variables. From Fleishman & Hempel (1956).

Elsewhere, Jones has suggested that we start with the learning trial correlations, analyze them first, and then estimate the loadings of reference test measures on the factors defined only by the learning trials. I have felt the need to start with the reference tests; my own feeling is that we know a great deal more about these tests. In fact, much of our work has centered on trying to get better definitions of these tests as a prelude to their use in the study of skill acquisition questions. However, I have tried it the way in which Jones has suggested. What happens most often is that one gets three learning trial factors with one factor decreasing as a function of practice. This is the factor that tends to have the reference tests loaded on it. But their differential contribution is confounded in this single general "pretask experience" factor. A second factor is found which increases in contribution with practice; this is interpretable in terms of our "within task-specific" variance, which is always shown to increase in contribution with practice. Jones, in fact, found this to be the case in his reanalysis of one of my studies (Jones, 1962). My own feeling is that the studies with most psychological meaning have been those in which the learning trials are interpreted in terms of the more general reference ability factors.

Jones emphasizes the great deal of specificity associated with motor performance and this cannot be explained away by our results. In fact, our work underscores his assertion. Performance does appear to be increasingly a function of habits and skills acquired in the task itself. But abilities play a role too, and their interactions with learning phenomena are important sources of variance to be studied. Furthermore, it appears desirable to better define the variance now termed specific to individual tasks. I am optimistic that some of this variance is not really "specific"; rather, we may need to be more ingenious at teasing it out.

Much of our later work has been concerned with a pursuit of the variance now defined as specific to late stages of learning tasks. Hypotheses we have explored are that (a) late stage measures of different practiced tasks have abilities in common not found in early stages of the same tasks (Fleishman, 1957), (b) the ability to integrate component abilities represents a separate individual difference variable not found in early stage learning, but is critical in late stage learning, (c) kinesthetic ability factors play an increasing role in psychomotor learning relative to spatial-visual abilities. Figure 4 shows the confirmation of the latter hypothesis in a recent study (Fleishman & Rich, 1963). In this last study it is useful to point out that the exploitation of individual differences allowed us to test a principle of motor learning which had only been assumed before (Fitts, 1951). Our findings on individual differences and skill acquisition are consistent with an information processing model of learning, where abilities are seen as capacities for processing different kinds of information. [For an elaboration, see Fleishman & Rich (1963).]

Lately our findings and methods have been extended to more complex tasks studied over lengthy periods of time. In one study (Parker & Fleishman, 1960) we developed a simulation of an air intercept mission on which subjects performed over a 7-week period and we have studied the acquisition of skill in a real training environment (Fleishman & Fruchter, 1960). In the latter case, we were able to show the abilities underlying the acquisition of skill, at different stages of Morse code learning, in radio telegraphy. We have also shown how knowledge of the abilities underlying performance at particular

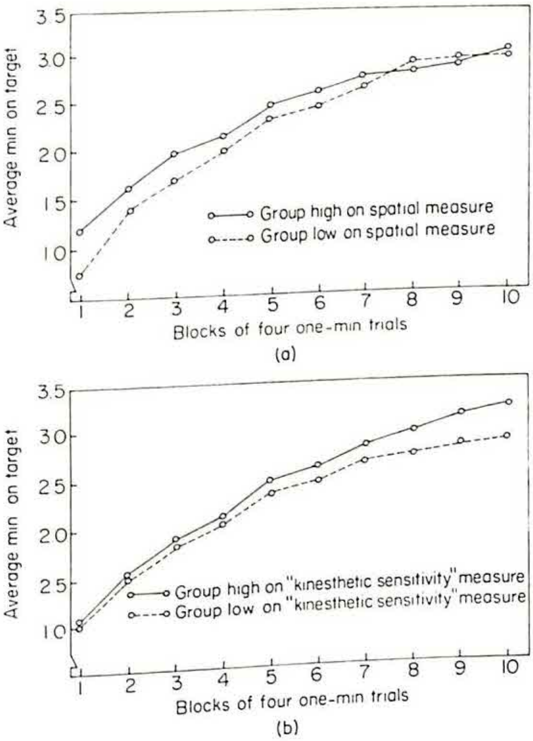


FIG. 4. Comparison of Two Hand Coordination acquisition curves for groups high and low on two different abilities. From Fleishman & Rich (1963).

stages of learning can be used to facilitate the training of a complex skill (Parker & Fleishman, 1961).

## Individual Differences in Retention

I would not like to leave the topic of individual differences and skill acquisition without mentioning our studies on retention (Fleishman & Parker, 1962). Very little is known about individual differences in retention. We were able to give people extended (7 weeks) practice on a highly complex perceptual-motor skill and obtain subjects back after periods of no practice of 1, 4, 9, 14, and 24 months. Thus, we varied retention intervals, as well as type of initial guidance and level of original learning. The main points of interest here are that there was virtually no loss in skill regardless of the length of the retention interval, and that the most powerful variable operating was individual differences in the level of original learning. The prediction of retention from original learning was independent of the length of the retention interval. Thus, for all intervals, even up to 2 years, individual differences at the end of learning correlated in the .80's and .90's with subsequent performance after periods of no practice. Our design also allowed us to say that this prediction was not accounted for by the subject's pretask abilities, but rather was explainable in terms of individual differences among subjects in the specific habits acquired in practicing the original task.

## Individual Differences and Other Learning Phenomena

We have also investigated the relation between individual differences and variety of learning phenomena, including associative interference and transfer, reminiscence, and performance during massed vs. distributed trials (Fleishman & Ellison, unpublished work, 1961). Space does not permit us to summarize these studies here, except to say that knowledge about subjects' previously developed abilities helped us to predict these phenomena. We also introduced "personality" variables here, and these were notably unsuccessful in predicting interference and transfer phenomena. However, they were more successful than ability variables in predicting the performance decrement which occurred during our massed practice period.

A recent study (Fleishman, 1965) investigated the relations between individual differences in performance on task components and subsequent performance on a total task (a multidimensional compensatory tracking task). These results pointed up the oversimplification of explanations based on simple S-R concepts. The level of description which provided the "best" explanation of the observed relations among the components and with the total task were in terms of "common ability" requirements.

## Abilities and Task Requirements

In several studies (e.g., Fleishman, 1956) we have been able to show systematic changes in the abilities required to perform certain tasks, as the display-control relations in these tasks were systematically varied. In all cases subjects received a battery of reference measures prior to experiencing the task manipulations. It was found that progressive rotations of the display panel changed the ability requirements of the task from "perceptual speed" to two other factors, "spatial orientation" and "response orientation." Thus, individual differences along known dimensions are used to explore the relations between tasks and the characteristics of people who could perform the tasks most effectively. Of course, these are problems faced every day by the personnel, engineering, and systems psychologists. The significance of the above studies is that a body of principles about these relations may be built up in the laboratory using the ability-skill paradigm which we described earlier.

## Concluding Statement

From all this it seems to me that the fact of individual differences can be exploited to gain insights into the mechanisms of skill learning. The subjects' pretask abilities become major treatment variables with significant interactions with learning trials and with other learning phenomena. At the beginning of his paper Jones quoted Cronbach's (1957) lament that we have two scientific disciplines of psychology. But Jones is still pessimistic that this division can be bridged; rather he sees more hope in parallel lines of investigation. At the time of his APA presidential address, Cronbach (1957) quoted our earlier work as hopeful examples of attempts to bridge this gap. I am not so sure to what extent this prophecy has been fulfilled. As Jones has pointed out, there are still many unresolved conceptual and methodological problems. But we feel there is much to be gained in the study of individual differences and learning within the same conceptual framework, through a combination of correlational and experimental methods; and we wish more psychologists, particularly "learning psychologists," would join in the quest.

### References

Adams, J. A. The prediction of performance at advanced stages of training on a complex psychomotor task. *USAF Hum. Resources Res. Center res. Bull.*, 1953, No. 53–49.

Bechtoldt, H. P. Factor analysis and the investigation of hypotheses. *Percept. mot. Skills*, 1962, **14**, 319–342.

Brown, J. S., & Jenkins, W. O. An analysis of human motor abilities related to the design of equipment and a suggested program of research. In P. M. Fitts (Ed.), *Psychological research on equipment design.* AAF Aviation Psychol. Program Res. Rep. No. 19. Washington, D.C.: U. S. Govt. Print. Off., 1947.

Cronbach, L. J. The two disciplines of scientific psychology. *Amer. Psychologist*, 1957, **12**, 671–684.

Ferguson, G. A. On learning and human ability. *Canad. J. Psychol.*, 1954, 8, 95–112.

Ferguson, G. A. On transfer and the abilities of man. *Canad. J. Psychol.*, 1956, **10**, 121–131.

Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 1287–1340.

Fleishman, E. A. A factor analysis of intra-task performance on two psychomotor tests. *Psychometrika*, 1953, 18, No. 1. (a)

Fleishman, E. A. Testing for psychomotor abilities by means of apparatus tests. *Psychol. Bull.*, 1953, **50**, 241–262. (b)

Fleishman, E. A. Dimensional analysis of psychomotor abilities. *J. exp. Psychol.*, 1954, **48**, 437–454.

Fleishman, E. A. Psychomotor selection tests: research and application in the United States Air Force. *Personnel Psychol.*, 1956, **9**, 449–467.

Fleishman, E. A. A comparative study of aptitude patterns in unskilled and skilled psychomotor performances. *J. appl. Psychol.*, 1957, **41**, 263–272. (a)

Fleishman, E. A. Factor structure in relation to task difficulty in psychomotor performance. *Educ. psychol. Measmt.*, 1957, **17**, 522–532. (b)

Fleishman, E. A. An analysis of positioning movements and static reactions. *J. exp. Psychol.*, 1958, **55**, 13–24. (a)

Fleishman, E. A. Dimensional analysis of movement reactions. *J. exp. Psychol.*, 1958, **55**, 438–453. (b)

Fleishman, E. A. Abilities at different stages of practice in Rotary Pursuit performance. *J. exp. Psychol.*, 1960, **60**, 162–171. (a)

Fleishman, E. A. Psychomotor tests in drug research. In J. G. Miller and L. Uhr (Eds.), *Drugs and behavior*. New York: Wiley, 1960. (b)

Fleishman, E. A. The description and prediction of perceptual-motor skill learning. In R. Glaser (Ed.), *Training research and education*. Pittsburgh, Pa.: Univer. of Pittsburgh Press, 1962.

Fleishman, E. A. Factor analyses of physical fitness tests. *Educ. psychol. Measmt.*, 1963, **23**, 647–661.

Fleishman, E. A. *The structure and measurement of physical fitness.* Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Fleishman, E. A. The prediction of total task performance from prior practice on task components. *Hum. Factors*, 1965, **7**, 18–27.

Fleishman, E. A., & Ellison, G. D. A factor analysis of fine manipulative performance. *J. appl. Psychol.*, 1962, **46**, 96–105.

Fleishman, E. A., & Fruchter, B. Factor structure and predictability of successive stages of learning Morse code. *J. appl. Psychol.*, 1960, **44**, 97–101.

Fleishman, E. A., & Fruchter, B. Component and total task relations at different stages of learning a complex tracking task. *Percept. mot. skills*, 1965, **20**, 1305–1311.

Fleishman, E. A., & Hempel, W. E., Jr. Changes in factor structure of a complex psychomotor test as a function of practice. *Psychometrika*, 1954, **18**, 239–252. (a)

Fleishman, E. A., & Hempel, W. E., Jr. A factor analysis of dexterity tests. *Personnel Psychol.*, 1954, **7**, 15–32. (b)

Fleishman, E. A., & Hempel, W. E., Jr. The relation between abilities and improvement with practice in a visual discrimination reaction task. *J. exp. Psychol.*, 1955, **49**, 301–312.

Fleishman, E. A., & Hempel, W. E., Jr. Factorial analysis of complex psychomotor performance and related skills, *J. appl. Psychol.*, 1956, **40**, 96–104.

Fleishman, E. A., Kremer, E. J., & Shoup, G. W. *The dimensions of physical fitness—a factor analysis of strength tests.* ONR, Contract Nonr 609(32), Tech. Rep. 2, Yale Univ., 1961.

Fleishman, E. A., & Ornstein, G. N. An analysis of pilot flying performance in terms of component abilities. *J. appl. Psychol.*, 1960, **44**, 146–155.

Fleishman, E. A., & Parker, J. F. Factors in the retention and relearning of perceptual-motor skill. *J. exp. Psychol.*, 1962, **64**, 215–226.

Fleishman, E. A., & Rich, S. Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. *J. exp. Psychol.*, 1963, **66**, 6–11.

Fleishman, E. A., Thomas P., & Munroe, P. *The dimensions of physical fitness—a factor analysis of speed, flexibility, balance, and coordination.* ONR, Contract Nonr 609(32), Tech. Rep. 3, Yale Univ., 1961.

Gagné, R. M., & Fleishman, E. A. *Psychology and human performance: an introduction to psychology.* New York: Holt, 1959.

Greene, E. B. An analysis of random and systematic changes with practice. *Psychometrika*, 1943, **8**, 37–53.

Hempel, W. E., Jr., & Fleishman, E. A. A factor analysis of physical proficiency and manipulative skill. *J. appl. Psychol.*, 1955, **39**, 12–16.

Herbert, M. Analysis of complex skill: vehicle driving. *Hum. Factors*, 1963, **5**, 363–372.

Jones, M. B. Practice as a process of simplification. *Psychol. Rev.*, 1962, **69**, 274–294.

Locke, E. A., Zavala, A., & Fleishman, E. A. Studies of helicopter pilot performance: I. The analysis of task dimensions. *Hum. Factors*, 1965 (in press).

Melton, A. W. (Ed.). *Apparatus tests.* Washington, D.C.: U.S. Govt. Printing Office, 1947. AAF Aviat. Psychol. Prog. Res. Rep. No. 4.

Nicks, D. C., & Fleishman, E. A. What do physical fitness tests measure? A review of factor analytic studies. *Educ. psychol. Measmt.*, 1962, **22**, 77–95.

Parker, J. F., & Fleishman, E. A. Ability factors and component performance measures as predictors of complex tracking behavior. *Psychol. Monogr.*, 1960, **74**, No. 16 (Whole No. 503).

Parker, J. F., & Fleishman, E. A. Use of analytical information concerning task requirements to increase the effectiveness of skill training. *J. appl. Psychol.*, 1961, **45**, 295–302.

Reynolds, B. The effect of learning on the predictability of psychomotor performance. *J. exp. Psychol.*, 1952, **44**, 189–198.

# Chapter 4

# Some Mechanisms of Motor Responding: An Examination of Attention

JACK A. ADAMS
*University of Illinois*

This chapter deals with a subject that some might consider close to perception, but with the emphasis in this volume on perceptual-motor behavior, I do not think the focus strained. Primarily the discussion is directed toward attention processes. *Attention* is a term that has not entered the technical vocabulary of many experimental psychologists but, nevertheless, the gist of this paper is to urge closer examination of behavioral phenomena that might fall under the heading of attention. As an introduction, the background of attention and some methodological issues are sketched.

It is instructive to recall the prebehaviorist period of psychology when it was mandatory for any psychological system worthy of the name to give central emphasis to attention. The method of that early psychology was introspection of the properties of one's own consciousness, with the goal of finding basic mental elements and how they were lawfully organized as the complexities of aware, conscious states. With private consciousness as the subject matter, it is little wonder that attention drew such interest. One need not be a trained introspectionist to know that his eyes move from one object to another as first one thing and then another is fixated, that our awareness is subjectively clearer at the center of the visual field than at the periphery, that the number of items we can handle in awareness at any moment is limited, that attentiveness is accompanied by a preparatory muscular set, that certain classes of stimuli intrude on our awareness more readily than others, and so on. If description of consciousness is the scientific goal of one's psychology, it is apparent that one cannot ignore attention as a dominant phenomenon. Titchener, the dean of American experimental psychologists of that time, in his "Lectures on the Elementary Psychology of Feeling and Attention" published in 1908, went so far as to say that ". . . the doctrine of attention is the nerve of the whole psychological system, and that as men judge it, so shall they be judged before the general tribunal of psychology" (Titchener, 1908, p. 173). This was the time that entire books were written about attention and all textbooks of psychology had a major chapter on it.

The decline of interest in attention in the United States was dramatic. While any intellectual movement has amorphous beginnings, it seems fair to date the strong beginning of behaviorism from Watson's paper (1913), "Psychology as the Behaviorist Views It," and attention began to fade about that time also. Wundt (see Boring, 1929, p. 327) distinguished between *mediate experience* whereby the physical scientist passively observed relations in the world without attending to consciousness per se, and *immediate experience* where the psychologist attended the conscious representations of the real world by turning his introspective mind's eye inward. Thus, psychology had a subject matter and a method that were unique among the sciences, and it was this unquestioned first principle of psychology that behaviorists shattered. Watson publicized the assertion, which is now commonplace, that *all* sciences have a common method which is indifferent to consciousness per se, and that the psychologist's task is to manipulate, objectively observe, and measure behavior just as a physicist treats phenomena of the physical world. A behaving organism is just another system of the physical world whose laws are to be determined. This psychology caught hold and its successes in terms of objective behavioral laws came early and were impressive. And, with the adoption of behaviorism, attention as a giant of consciousness-centered psychology quickly faded. By the early 1940's the concern with attention was minor but it had not died and Paschal (1941), in a review, was slightly encouraged to find a low-level, persisting interest in attention throughout psychology. There was an unwillingness to let attention die, and Paschal wrote: ". . . despite the criticisms that have been aimed at it, the attempts to read it out of the party, attention has had a heartiness that has enabled it to survive" (Paschal, 1941, p. 383). Looking at attention today, one must agree with Paschal on its capacity for survival.

Today we are riding a new wave of interest in attention. This wave was accelerated by two forces: *first*, an interest in vigilance that had its main beginnings in World War II, and *second*, physiological research on the brain stem reticular formation that began in earnest about 1949. Psychology and physiology journals of recent years have hundreds of articles where attention is a term liberally sprinkled about, and it is fair to say that attention has staged a comeback. Let me briefly mention vigilance research as a force in the attention revival. During World War II radar was an exciting technical development that proved decisive in many air battles. However, a source of system degradation was found to be air surveillance operators who sat at radar scopes hour after hour, watching a monotonous display for the appearance of enemy aircraft. They sometimes missed radar returns that should have been obvious and detected quickly. Laboratory research established that the human operator's capability for detecting these critical events decreased over a long session of continuous watchkeeping and, since the task of watching was not physically demanding, performance decrement was ascribed to the decline of long-term attention, or vigilance, rather than muscular fatigue. Interest began to develop in the determinants of vigilance, but it was not until the publication of N. H. Mackworth's important monograph in 1950 (see Mackworth, 1961) that programs of research and theorizing got underway in a lively fashion.

The second force for our modern interest in attention was research on the brain stem reticular formation, or the "old brain" (Delafresnaye, 1954; Jasper, 1958; Lindsley, 1957; Magoun, 1963; Rossi & Zanchetti, 1957; Samuels, 1959). The classical view of the afferent-efferent arc in the central nervous system was an afferent input that went to its center in the sensory cortex and made connection with an efferent system whose activation was revealed in an overt response. Essentially beginning with the research of Moruzzi and Magoun (1949), and in a large number of studies since then, this classical view has been greatly elaborated. Not only are there rather direct afferent paths to the sensory cortex but, in addition, there are collateral paths to the reticular formation. Activation of the reticular formation through secondary paths produces a massive electrical effect upward to the cortex which is observed as desynchronization of the EEG (electroencephalogram), or the change of electrical potential of the cortex from relatively slow wave forms with reasonably large amplitude to fast waves of low amplitude. When this activation, or arousal as it is often called, accompanies a stimulus the organism will customarily show all the behavioral signs of being awake, alert, and attentive. In the absence of arousal the organism will usually be behaviorally indifferent to stimulation or even asleep. The response of an organism is now thought to depend *both* on an evoked potential at the sensory cortex and firing of the reticular formation. A stimulus not only has a cue function but an arousal function too, and response to a stimulus depends on both functions being present.

On the behavioral side, work on the reticular formation provoked two main lines of thinking. One line, which seems most prominent among physiologists, reasons that if the organism is alert and attentive when the reticular formation is aroused, and indifferent or asleep when the reticular formation is quiescent, then attention is physiologically explained by the reticular formation. Using the reticular formation in this way is satisfying to some from two points of view. First, when the reticular formation of a quiescent or sleeping animal is stimulated he assumes postural alertness, pricks up his ears, and looks around in a manner that informally we all have come to recognize as attentive. Second, EEG research on humans has shown over and over, in many different ways, that desynchronization of EEG occurs when there is orientation to a distinctive stimulus, when a problem is being solved, or sometimes even when one is thinking about stimuli or problem solving, and again we can agree that this informally meets the layman's criterion of paying attention.

A second main line of thinking is centered more in psychology and suggests that the reticular formation is the basic center of generalized motivation. Hebb (1955) and Malmo (1959) have been the leading advocates of this interpretation. The *general* property of the motivational state should be emphasized, and it can be considered similar to Hull's drive concept (Hull, 1943) in having the power to energize all response classes. Malmo (1959) bolsters this assertion with evidence that a number of psychophysiological measures such as EEG, electromyogram, heart rate, and galvanic skin response covary under conditions which psychologists usually consider motivating, and he reasons that the reticular formation is the central agent that affects all these measures. An important implication of this position is that superthreshold stimuli have a motivating charge associated with their occurrence, or at least until habituation sets in (Sharpless & Jasper, 1956). Since the reticular formation is a center for probably all stimulus afferents both of peripheral and central origin, and since the reticular formation is hypothesized to be the seat of motivation, it follows that all stimuli are motivating. Traditional sources of motivation such as sex might operate through hormonal influences on the reticular formation, and more derived forms of social and secondary motivation would work upon the reticular formation from higher cortical centers downward. All of this is an expanded view of motivation, particularly the position that all stimuli are intrinsically motivating. Too, this research is important for building an understanding of the physiological bases of motivation. However, it does not help the problem of defining attention. At most, it says that the behavior we decide to classify as attentive can be heightened by motivation, which is scarcely surprising. In some of my own studies of vigilance (Adams & Boulter, 1962; Adams, Stenson, & Humes, 1961), I have found heuristic value in this notion that stimuli and their properties can be motivating and enhancing for human performance.

Despite the importance of this research and these lines of thinking, I find it regrettable that physiological research on the reticular formation, with its emphasis on general alerting, ever became associated with attention, because I think (and this is *my* definition) that *attention is best used as a descriptive term for directive, discriminative response to a stimulus subset of a general* *stimulus field*. In this sense I find a measure of agreement with Oswald (1962), who holds that attention is always *to* something, and with physiological psychologists who speak of *specific* alerting mechanisms (e.g. Lindsley, 1957, p. 87). During attention's finest hour the introspectionists held a similar view. It was only with the advent of research on the reticular formation that attention became associated with general alerting. General alerting, or motivation, as some have come to think of this action of the reticular formation, is certainly an element of laws describing response to a stimulus subset, but I am disinclined to equate behavioral arousal and attention regardless of how much the organism pricks up his ears and looks around bright-eyed. Specific alerting, with all the attendant mechanisms for making a response to a stimulus subset, is the heart of attention from my point of view.

In making this definition of attention it should be clear at the outset that I am not arguing for a special concept of attention as a homunculus—the little man inside the head who determines a stimulus subset that will receive response. A homunculus concept that presumably would take its place in theory alongside habit, motivation, and the like, seems to smack a bit of faculty psychology where an entity of the mind is postulated which conveniently has all the directing properties for the class of behavior being explained, and all without operational definition. I am not at all convinced that we need a scientific concept of attention as such, although our data will force it if it is needed, but we do need an increased concern with variables that determine how a particular stimulus subset becomes an elicitor for response. Underwood (1963), in a recent analysis, has shown concern with this matter of selective attending in verbal learning and he finds no particular need for a concept of attention. Rather, his approach shows the topic amenable to S-R (stimulus-response) learning analysis. Actually, an S-R approach to selective attending is not new. Estes (1959) dealt with this matter in his statistical learning theory when he postulated random sampling of stimulus subsets from trial to trial, and both Guthrie (1959, pp. 178–179) and Underwood (1963) distinguish between nominal stimuli of the physical world which abundantly impinge on the subject and the cue or functional stimulus to which the subject responds. Guthrie said that the change of nominal stimuli to cues is a result of learning, and this is hardly disputable. The whole study of learning can be construed as the study of attention because a part of the stimulus complex comes eventually to evoke a discriminative response, but this would hardly be a useful conceptualization. What would be useful, however, is a better grasp of how stimulus subsets from the great welter of those impinging become effective for response, and how these selective acts pattern themselves over time and change with learning, motivation, work inhibition, etc. Many lean on the legacies of Guthrie and Hull that learning occurs to all stimuli impinging at the moment of response, but it is becoming increasingly clear that the human being is not a *tabula rasa* on which all stimuli write equally, and that the stimuli which the experimenter presents and manipulates for response may not enter totally into the S-R bond on each trial. The subject may attend only to aspects of the stimulus complex, and sort out some for response at one time and others for response at another. Waugh (1961) has an interesting study of the free recall of serial lists where she presents evidence to show that a subject on a trial accepts only enough items to fill the capacity of immediate memory and learns only them while ignoring the others. On the next trial another subset of items is selected, and so on until the list is learned. The subject uses a piecemeal strategy to learn the task in a manner quite different from assumptions of older S-R models of verbal learning that assume equal attention to all items and an increment of habit strength for each item on a trial. Here our understanding of attention, of how the subject comes to choose a stimulus subset, involves the subject's strategy in attacking a memory task of heavy load and the characteristics of items that urge their selection on a trial. When we come to grapple more fully with matters such as these I think we will be regaining touch with the subtle, selective behavior that the introspectionists call attention, and I consider this a robust trend because in their own way I believe they were in contact with important behavioral phenomena that need lawful explanation.

Our modern approaches toward behavioral observation and measurement are crude expressions of attentiveness as I have defined it, and they gloss over most or all of the behavioral details which are at the heart of attention as psychologists have conceived it for almost a hundred years. Most of our experiments use such behavioral measures as percentage correct, number of responses per unit of time, response latency, time on target, trials to criterion, integrated error, etc. All of these measures grossly reflect attentive behavior, because obviously stimuli are received, but we have ridden roughly over the fine grain of behavior that once was attention. There is no methodological issue associated with these gross behavioral units because lawfulness can proceed at any level of description. Nevertheless, these big units conceal the fine fabric of behavior where attention is woven, and I think that neglect of behavioral detail is attenuating the goal of good predictive laws even at the level of large-scale description where experimental psychology usually chooses to operate. I would suggest that a goodly amount of variability occurs in behavioral measures because subtle behavioral details operate in different amounts at different times during the course of our observations. For example, a subject may or may not be visually oriented to a stimulus when it occurs, and he may or may not respond or he may respond slowly. Or he may or may not have a suitable anticipatory state and suitable muscular readiness. Increases in variance certainly will result. Theorists like Hull and Estes built mechanisms of variability into their theories, undoubtedly because the details of the responding and their determinants were ignored and they felt obligated to give some account of the ubiquitous variability of behavioral data. I am not so optimistic to think that variability will ever be completely removed from behavioral measures, but we must prod ourselves from time to time and remember that much of variability comes from the action of disregarded variables and mechanisms, and the precision of our behavioral laws is dependent upon knowing them. And, by minding the details of responding, we might hope to describe them lawfully and to predict, say, which element of a display will be selected and given motor response for how long until change is made to another element, or which nonsense syllable will be chosen for memorization on a trial. Our present behavioral measures cloud these details because they contain a myriad of them, but prediction of these response details strikes me as an admirable scientific goal.

These general remarks on attention lay the groundwork for a swing back to motor behavior and what attention means for it. I would like to examine research on three prominent mechanisms that determine at any moment the stimuli that elicit a response: the observing response, anticipation, and human capacity for processing multiple stimulus streams. The empirical findings in behalf of these mechanisms often have come from measures of massive chunks of responding, but they nevertheless bear on the subtle flow of attentive behavior.

## The Observing Response

Tasks used in the study of motor behavior often have relevant stimuli varying spatially from moment to moment. In a one-dimensional visual tracking task the subject follows the display events with his eyes in order to receive the data that command his response. Our venerable friend the Rotary Pursuit Test is a good example of an active requirement for visual responding that importantly determines the accuracy of the criterion motor response. In multidimensional visual tracking the burden for eye movements, and sometimes head movements, is doubly heavy because the subject must select the stimulus source for his response of the moment as well as visually follow the events of the source for a time until shift is made to another source. Complex displays such as these, which are common for motor behavior, represent instances where the subject is physically unable to receive all stimuli at once, and our assignment is discovery of the variables and mechanisms that determine stimulus reception.

We have, then, as a basic element for attention in visual-motor tasks, a small-scale response for the visual system that intervenes between the display and the motor response that we measure. It is a response chain that we are dealing with—an observing response as antecedent to the motor response. Wyckoff (1952) held that an observing response is any response that determines stimulus reception, and visual receptors are the most prominent example of it although, by Wyckoff's definition, other responses can function for stimulus reception as well. In any event, occurrence of the criterion motor response is often dependent upon the visual observing response and study of stimulus properties that evoke it, as well as its learning and retention, should give us an improved capability for predicting motor behavior. Certainly a portion of the variance in motor-learning curves must stem from learning the visual observing response, just as motor-forgetting functions in part probably represent loss in effectiveness for the observing response. Irion (1948) included changes in receptor orienting mechanisms as a contributor to loss of set which is revealed as warm-up decrement. No one has explored this feature of Irion's hypothesis, but his thinking is consistent with the thesis that greater concern with fine-grain behavioral details is a valuable part of behavioral lawfulness.

### *The Observing Response and Measurement of Eye Movements*

The direct measurement of eye movements has most often been a laboratory method carried under the rubric of perception, but hardly ever motor skills. Basic perception has been concerned with how complex forms are scanned and assimilated, and applied studies of reading and advertising have had similar interests. Ford, White, and Lichtenstein (1959, p. 287) list four choices of research vehicles for collecting eye movement data. First, the pure optical procedure by which eye movements are plotted from a light beam reflected from a small mirror attached near the cornea and registering in two dimensions on photographic paper. Second, eye camera procedures whereby reflected light from the convex corneal surface is photographed with a camera using continuous moving film (the ophthalmograph; e.g. Arnoult, 1958). Third, analysis of motion pictures of the eyes during execution of the visual task (e.g. Baker, 1960; Jones, Milton, & Fitts, 1949). And fourth, electro-oculography by which a plot of eyeball movements is made from a pickup of the corneo-retinal potential (e.g. Ford, 1957; Ford *et al.*, 1959; White & Ford, 1960).

J. F. Mackworth and Mackworth (1958) comment that a difficulty with these various methods is relating eye fixations to positions on the visual scene, and their ingenious solution combines the corneal reflection technique with closed circuit television. One television camera scans the pip of light reflected from the corneal surface, another scans the visual scene, and the two pictures are electronically mixed to give a continuous showing of the scene with a moving pip of light always showing where the subject's eye is fixated. The apparatus is cumbersome and requires a bite bar for keeping the head rigidly in place, but recent improvements by Mackworth and his associates substitute a small motion picture camera for television equipment and the whole assembly is mounted on a helmet which the subject wears. This arrangement not only gives mobility to the subject, but also gives a permanent film record which investigators require. This recent development would seem the most promising of all for the study of eye movements and skills, but no studies have come to my attention as yet.

The principal difficulty in using eye movements for inferring observing responses is that they measure looking, not seeing. Thus, they fail as a primary measure of an observing response which functions to provide the discriminating stimuli for the criterion response, which is a motor response for our purposes here. Having the eyes properly oriented is part of the battle, but if we are interested in how stimuli of motor tasks are received then we must remember that stimulation capable of exciting the motor responses is far more complex than looking. For example, the subject has some capability for receiving stimuli in peripheral vision even though he is not looking. There is the matter of visual threshold which can deny seeing even though the subject is looking. The visual receptors can be "fatigued" and have a reduced capability for seeing even though the subject is looking. Studies of vigilance by Bakan (1955), McFarland, Holway, and Hurvich (1942), and Zuercher (1965) have shown that with prolonged use the visual system becomes increasingly less sensitive. In a study using the Rotary Pursuit Test, Adams (1955) found that a period of watching a partner practice and discriminating his on-off target performance was sufficiently fatiguing for the visual system to lower the subject's arm-hand performance in a subsequent practice session.

Because of these various differences between looking and seeing—between being merely oriented to a stimulus and discriminating it—I think that we must throw our weight in the direction of seeing or, more generally, the observing response, which not only implies receptor orientation but also stimulus reception and discrimination for the criterion response that follows. If this point of view is granted, it indicates for this state of the scientific game that we have an emphasis on molar behavior and use the criterion motor response for inferring the observing response. Sanders (1963) holds a similar view. Measures of eye movements and looking should not be totally discounted and set aside, however. Measurement of eye movements can be worthy if for no other reason than that motor behavior and eye movements have not been studied jointly, and such research may bear far more fruits than my remarks here have indicated. Certainly measures of eye movements could be expected to account for part of the variance of motor behavior. Nevertheless, I stand by my general criticism that measures of eye movements are at best a half-way house and that inferences of the observing response from measures of the criterion motor response itself are probably best. This approach is old hat for experimental psychologists. Frequently we infer about remote behavioral states from the gross, large-scale behavioral patterns that are our dependent variables. Inferring the internal fear or anxiety response from overt behavior is an example.

### *Inferring the Observing Behavior from Motor Responses*

Let me discuss studies that cast light on workings of the observing response as inferred from motor responses. A basic variable for the observing response is number of stimulus sources to which a subject must attend. If the subject's task is to find a particular stimulus quantity or quality from a number of stimulus sources, the greater the number of stimulus sources the greater the average search time to locate and report the value. In a nonmotor study Senders (1952) used a tachistoscope and studied the effects of number of dials on reading time and reading errors. The number of dials studied varied from 1 to 45. Her results showed that both reading time and reading errors were a linear increasing function of number of dials displayed. The greater the labors assigned the visual scanning mechanism the longer the time taken to detect a critical event and the poorer the performance.

A closely related variable is spatial separation of stimulus sources, and this has been most fully investigated with multidimensional visual tracking tasks, although Sanders (1963) made a noteworthy study that used keypressing as a response. It is reasonable to assume that the greater the separation between sources the longer the time taken for the observing response to travel back and forth and the greater the likelihood that tracking error would build up at a source. Empirical data support this expectation. Fitts and Simon (1952) used a two-dimensional compensatory pursuit task and found that time-on-target scores decreased as separation of two stimulus sources was increased from 4 to 16 inches. Similarly, King (1961) used a two-dimensional compensatory pursuit tracking task and found that time on target decreased as separation of the two sources was increased from 3.5 to 10 inches. A comprehensive study was described by Briggs and Howell (1959), who used a two-dimensional compensatory tracking task and varied the separation of stimulus sources from 2.1 to 10.7 inches. Frequency for a sine wave input was also varied and had values of 4, 8, or 12 cycles per minute. Consistent with the other studies, tracking error increased as separation of sources increased. The deterioration of tracking proficiency as a function of spatial separation was about the same for all input frequencies. All studies of tracking show that spatial separation of stimulus sources, as a determiner of the observing response, is a potent variable for motor performance.

In the case of multidimensional tracking where stimuli are dynamic and change with time, the subject has error building up on the other sources when he is correcting error at a given source. After error correction at a particular source he must then select another one. Assuming equal weight to stimulus sources, selection could be on the basis of a learned pattern for the observing response where the subject moves from one source to another in an orderly fashion, or it could be strategy of finding the source with the largest error and correcting it. When one or more sources are weighted more heavily than others as, for example, in landing an aircraft on instruments, there will be greater attention to those instruments which the pilot knows through his training to be most prominently related to success and safety. Or, if a subject is highly skilled and coordinated, he might scan two or more sources, discern error, and then correct them all together with a set of unified movements. We know precious little about shifts in attention for complex panel arrays, and even less about coordination and how a subject might achieve the high skill of apprehending events of two or more sources and giving them a simultaneous, coordinated response.

The studies discussed so far allow inferences about workings of the visual observing response from manipulation of task variables that are assumed to affect it, and this is a useful approach. However, a research project of mine went a step beyond and, with a fine-grain analysis of response records in a two-dimensional visual tracking task, made direct measurement and inferences about the observing response, its patterning, and what it meant for motor proficiency in the overall task. The equipment was a pursuit tracking task that had an analog-digital recording system where the subject's tracking error values were recorded in digital form for analysis by a computer. In an initial study (Adams & Webber, 1961) digital records of tracking error were printed out and it was noticed that they were typified by *error peaks:* where error increased, the subject sensed it, and then eliminated it with a corrective motor response. An error peak is a complex act of visual sensing and motor responding, but in a very real sense the observing response is the heart of it. The top of the peak, or the maximum point of the tracking error function between two successive segments of little or no error (on-target behavior), was taken as the point where a subject observed the error event and began response to null it. This initial study set our basic methodology, and our second study (Adams & Webber, 1963) fully explored ramifications of the error peak, including a computer model based on it that had provision for simulating one- or two-dimensional tracking, type of input signal, and practice trials. The details are rather elaborate, but for our purposes here I would like to mention some of the properties of error peaks in two-dimensional tracking and emphasize the value of fine-grain analyses. We developed a computer program for the automatic identification and measurement of error peaks from digital tracking error records. Furthermore, the program made zero-order probability, and first- and second-order conditional probability, analyses of peaks to document the patterns of responding in the task – that is, whether subjects dwelt repeatedly on a single stimulus source before shifting to the other, how they shifted from one source to the other, and how these observing response patterns changed with practice.

Table I is adapted from Adams and Webber (1963) and it shows error peak data from 12 subjects on two trials. An irregular input was used at each source $x$ and $y$. Forty practice trials were given, and Trials 4 and 40 are shown as representative of early and late practice. Presented in Table I are number of peaks for the group, zero-order probabilities of peak occurrence for $x$ and $y$,

TABLE I

ZERO-, FIRST-, AND SECOND-ORDER PROBABILITIES OF OCCURRENCES OF ERROR PEAKS IN THE $x$ AND $y$ DIMENSIONS OF A TWO-DIMENSIONAL PURSUIT TRACKING TASK WITH IRREGULAR INPUTS[a]

Trial 4

| | $P_i$ | *No. of peaks* | $P_{j(i)}$ | $x$ | $y$ | $P_{l(jk)}$ | $x$ | $y$ |
|---|---|---|---|---|---|---|---|---|
| $x$ | .57 | 663 | $x$ | .44 | .56 | $xx$ | .39 | .61 |
| $y$ | .43 | 489 | $y$ | .75 | .25 | $xy$ | .72 | .28 |
| | | | | | | $yx$ | .48 | .52 |
| | | | | | | $yy$ | .78 | .22 |

Trial 40

| | $P_i$ | *No. of peaks* | $P_{j(i)}$ | $x$ | $y$ | $P_{l(jk)}$ | $x$ | $y$ |
|---|---|---|---|---|---|---|---|---|
| $x$ | .55 | 669 | $x$ | .38 | .62 | $xx$ | .33 | .67 |
| $y$ | .45 | 542 | $y$ | .74 | .26 | $xy$ | .72 | .28 |
| | | | | | | $yx$ | .40 | .60 |
| | | | | | | $yy$ | .81 | .19 |

[a] The values shown are for the group of 12 subjects as a whole. Trial duration was 90 sec. From Adams and Webber (1963).

and the first- and second-order conditional probabilities for occurrences of peaks in $x$ and $y$. The center column has first-order conditional probabilities and gives, for example, the probability of a peak repeating in $x$, and the probability of a peak alternating to $y$ having just occurred at $x$. The right-hand column has second-order conditional probability which is the probability of occurrence in $x$ or $y$ given the immediate past history of the last two peaks. The most interesting thing about Table I is the surprisingly small change in number or patterning of peaks from early to late in practice, and it is evident that actions of observing responses are quite unchanging. Moreover, the computer program measured time between peaks on a trial and we found it did not change appreciably with practice. A subject's patterning of attention in the task—his selection of stimulus subsets for response—is somehow determined by the task demands themselves and has precious little to do with practice. One would be on unsure footing to conclude that the observing response is insensitive to learning, for subjects in some tasks must certainly develop more efficient modes of scanning, but it is surprising that error peaks show so little effects of learning. Other analyses disclosed that motor learning occurred in increasingly rapid correction of error once it had been detected, in more deliberate responding, and in the amount of time that a subject remained in the on-target zone once a correction was completed. I find these data support my contention that benefits are to be had in examining the fine grain of behavior, and that useful inferences about observing can be made from molar motor behavior.

Inferring the observing response from measures of the criterion motor response is not free of potential difficulties. An obvious problem to be examined is that inferences about the observing response contain variance from the motor response and the control system, so that if we held the display and visual requirements for the observing response constant we might conceivably infer different properties for the observing response. I do not know how serious this problem is, but it is accessible to research. By varying response requirements we should be able to understand how motor factors obscure evidences of the observing response. A good research tactic would be to start with simple display and motor response requirements and then augment the complexity to see how our inferences for the observing response are influenced. Another problem is that a subject may look and see but not make a motor response and so no record of his looking and seeing will be obtained. This may or may not be a problem area, and I see no way of clearly getting at it. Decision mechanisms seem to be involved. Despite these problems, I find the data which I have discussed a more informative means of understanding the observing response and variables associated with it than the measurement of eye movements.

## Anticipation

So far, implicitly, I have been mostly discussing the kind of motor task situation where the input is irregular and unpredictable, and where the subject has no advance information about task stimuli. A tracking task is often of this kind, where the subject must orient his visual observing response to the stimulus source, discern the requirements, and then execute the motor response. This is a classical paradigm in experimental psychology. A stimulus is presented, the subject senses it, makes an instrumental response, and as a result has response delay (off-stimulus time) that lags the stimulus by at least one reaction time value. As widespread as such tasks are, they fail to do justice to the range of possibilities for motor tasks and a major reason is that anticipation is absent in them. The pioneering thinking and research on anticipation was done by Poulton (1952a,b, 1957a,b,c), and he distinguishes between *receptor anticipation* and *perceptual anticipation.*

Receptor anticipation requires that stimulus events for future responding be available for the observing response in the current display. Receptor anticipation is simply preview of forthcoming stimulus events, whether they are random or regular, and the subject can ready and prepare himself for the immediate future. Looking ahead while driving and planning for a sharp curve is an example of receptor anticipation. Another example is a musician who sneaks brief glances ahead in the musical score to be optimally informed and ready for the moment when the notes are to be played. Short-term memory seems to be a mechanism here for retaining the advance information but little work has been done on mechanisms of receptor anticipation, and we cannot be sure that a musician's surreptitious glances benefit his performance. Leonard (1953) found preview advantageous for a discrete visual-motor task, but Gifford (1963) found that preview deteriorated response in pursuit tracking because the subjects were making anticipatory responses that proved damaging to ongoing performance. Whatever the influences of preview, it is a case where the subject is active in seeking advance information and remembering it. Poulton's label "receptor anticipation" admirably expresses the action of the observing response in casting ahead for stimuli of the future.

Perceptual anticipation, on the other hand, operates where there is no advance information in the outer world of the display. Through learning, the subject comes to know regularities in stimuli and can predict their occurrence with some accuracy. There are two kinds of perceptual anticipation: temporal and spatial. In temporal anticipation, the subject comes to predict when a stimulus event will occur, and in spatial anticipation the subject predicts where an event will occur. If a task has a single stimulus source the subject need only learn when stimuli will occur, but if there are multiple stimulus sources there can be learning of both when and where stimuli will occur. Perceptual anticipation obviously requires regularities in the patterning of events so there is a stability that can be learned, although we know little about the defining conditions for such regularities. The benefits for motor responding can be large because perceptual anticipation allows us to develop advantageous preparatory responses for the eventual motor response to the stimulus that is about to occur, and in the optimal case the motor response can be fired off slightly in advance of stimulus occurrence so the stimulus and response exactly coincide and error is zero. As the summit of achievement the subject could respond perfectly with his eyes shut—all his cues for responding are internalized. In effect, the observing response is now internal and the relevant stimulus subset for the criterion response becomes defined through learning.

Most of the research on anticipation has been with temporal anticipation, and a selective review would not be amiss for savoring the salient research directions. As research topics go in psychology, temporal anticipation has received relatively little work, and one should pause and reflect on this deficiency because timing is eminently the heart of so much skilled motor behavior. I have toyed with the possibility (Adams, 1964, p. 185) that our traditional concern with classical reaction time, with complete concentration on speed of response *after* the stimulus has occurred, has led us away from a forthright attack on anticipation. In classical reaction time studies, workers have always endeavored to avoid anticipatory tendencies and, as far as skills are concerned, we might well have been bypassing a big, potent mechanism.

Poulton's experiment (1952a) exemplifies the advantages that stem from temporal anticipation. Compensatory and pursuit tracking were compared for a simple harmonic input and a complex input composed of two sine waves. He was interested in anticipation of directional changes in the input, and he took as his measure the time difference between change in direction of the input and the corresponding change in direction for the response. For the regular harmonic input and pursuit tracking, where the subject had full information on characteristics of the input as well as feedback about his response to them, the response times after practice clearly showed the presence of anticipation because they were typically less than optimal classical reaction time (i.e., the value to a temporally unpredictable event where anticipation cannot be acquired). There was also anticipation for compensatory tracking with the simple harmonic course, but less than for pursuit tracking, and Poulton (1952a, b) concludes this was because compensatory tracking confounds information about the input and response to it, and these task conditions are poor for acquiring anticipation. With the complex input, gain in performance was due mainly to increased manual proficiency because anticipation was comparatively low and did not improve with practice. It is evident that anticipation underlies some of the differences we find between types of input signal and between compensatory and pursuit tracking.

Research in my laboratory has departed from Poulton's work and has been concerned with mechanisms of anticipation – that is, the underlying means by which a subject keeps time. Poulton's experiments, as well as traditional work in time perception (Woodrow, 1951), were at the descriptive empirical level and need to be elaborated with efforts toward mechanisms of timekeeping. In our first experiment (Adams & Creamer, 1962a) it was found that pretraining the neutral responses of pressing a button, saying a word, or moving a small crank in response to the directional change of a sine wave input transferred positively to anticipatory timing for the arm-hand following response when the same sine wave was used as input in a pursuit tracking task. We hypothesized that proprioception is a common type of internal stimulus associated with all these responses, and that fractional movements associated with the pretraining responses persist as the basis for anticipatory timing when transfer is made to the whole tracking task. Whether the subject has opportunity to form a mediation response through pretraining or has whole-task practice only, we hypothesized that the proprioceptive after-effects of response at time $t$ persist in time and provide internal cues to which the motor criterion response is learned for anticipatory occurrence at time $t + \Delta t$.

One could entertain a purely cognitive view where a visual perception of the temporal patterning of events is sufficient for acquiring anticipation, so a second study was run (Adams & Creamer, 1962b) where we manipulated task variables that are presumed to define the quantity and quality of proprioception (e.g., Bahrick, 1957) to see if proprioception could be related to to anticipatory timing. Using a pursuit tracking task again, but this time in a discrete configuration where a subject tracked step displacements of the input, we had a $2 \times 2 \times 2$ design where two values for each of spring tension, amplitude of movement, and signal duration were used. A subject had to learn to anticipate the time of occurrence of completely regular and repetitive step inputs, and we expected that greater amplitude of movement and spring tension would result in greater proprioceptive discharge and better timing. The longer signal duration should result in greater decay of the proprioceptive trace and poorer timing. Using as an index of anticipation for a subject the percentage of response times to input change that were less than optimal classical reaction time, we found some support for our hypothesis. Figure 1 gives the results. Both spring load and signal duration gave the expected effects but amplitude of movement did not, and we do not know why. Despite the failure for amplitude of movement, we concluded that proprioceptive stimuli appear to serve for response timing, in addition to their customary role of internal feedback that aids in the regulation of movements.

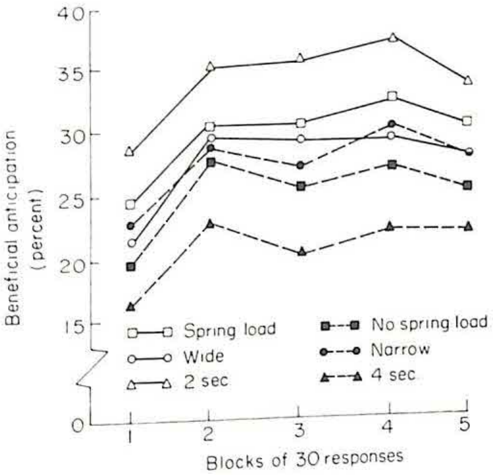


FIG. 1. Percentage of anticipation responses in one-dimensional discrete tracking (from Adams & Creamer, 1962b).

### *Anticipation in Complex Tasks*

In an experiment (Adams & Xhignesse, 1960) that was designed for the study of variables in two-dimensional tracking, it was found that temporal anticipation was a major factor in the efficiency with which subjects processed stimuli at two spatially separated sources. The tracking task has three linearly arrayed stimulus lights at each stimulus source and response was by a three-state control for each hand. The left-hand control was for response to the left stimulus source and the right-hand control was for response to the right-hand source. The task was fundamentally pursuit tracking because each stimulus light had a feedback light beneath it that was linked to a control, and the subject always had knowledge of both the stimulus event at a source and the position of his control. The response was to keep the feedback light aligned as much as possible beneath the stimulus lights as they changed. To provide for manipulation of the observing response, the two stimulus sources were on a semicircular track at eye level and each was placed 5 or 30 degrees on either side of the center line of vision, depending upon the experimental condition. Both a fast (one change per second) and a slow (one change per 2 seconds) stimulus rate were used.

Anticipation was possible through predictability of the center stimulus light at each source. In the high-coherency, predictable condition, the subject always knew that the center light would follow either of the other two lights after a fixed amount of time, so this was a case of both event and time certainty. A low-coherency, unpredictable condition had event certainty for the center light but temporal uncertainty where the center light always followed either of the other two lights with uncertain foreperiod and the subject was never sure when it would come on. The center light of the three in a source was white and the two lights adjacent to it were red and green. Predictability, or coherency, was defined by the conditional probabilities defining the sequencing of the three events. Table II shows that for a source in the high-coherency case no event follows itself and every colored event was always followed by a white event. However, a white event had a .50 probability of being followed by red or by green. Such a series for a source might look like red, white, red. white, green, white, red, white etc. In the low-coherency case, however, a colored event has a probability of .50 of being followed by itself or by a white event. This meant that a colored event, in sometimes repeating itself, persisted for longer than the 1 or 2 seconds defined as the duration of each event for any particular group, and the white event in effect had an uncertain foreperiod. Thus, with three successive red events for a 2-second condition, the subject would be confronted with a requirement in a source to hold the control aligned with red for 6 seconds. A

TABLE II
ZERO-ORDER AND CONDITIONAL PROBABILITY TABLES FOR THE TWO COHERENCY LEVELS OF THREE STIMULUS LIGHTS, RED, WHITE, AND GREEN AT EACH SOURCE[a]

High Coherency

| | | $P_{(i)}$ | | $P_{i(j)}$ | $j$ | | |
|---|---|---|---|---|---|---|---|
| | | | | | *Red* | *White* | *Green* |
| | Red | .25 | | Red | 0 | 1.0 | 0 |
| $i$ | White | .50 | $i$ | White | .50 | 0 | .50 |
| | Green | .25 | | Green | 0 | 1.0 | 0 |

Low Coherency

| | | $P_{(i)}$ | | $P_{i(j)}$ | $j$ | | |
|---|---|---|---|---|---|---|---|
| | | | | | *Red* | *White* | *Green* |
| | Red | .33 | | Red | .50 | .50 | 0 |
| $i$ | White | .33 | $i$ | White | .50 | 0 | .50 |
| | Green | .33 | | Green | 0 | .50 | .50 |

[a] From Adams and Xhignesse (1960).

low-coherency series in a source might look like white, red, red, white, green, green, white, red, red, red, etc. All events were treated as discrete and of fixed duration for the purposes of constructing the probabilistic, time-based series but allowing an event to repeat itself in the low-coherency case enabled the subject to see some events as longer than others. Thus, in the low-coherency case the within-source instability also gave low between-source coherency because the subject was never certain when a colored event would change to white in either source. This is to be contrasted with the high-coherency series where the within-source regularities give high between-source coherency because a colored event in the left source was always paired with a white event in the right-hand source, and both of them simultaneously changed to a white event in the left source and a colored event in the right source. In the high-coherency series the subject always had an event class present in one of the stimulus sources which he could potentially predict in time and type and reduce the requirement for visual scanning, while in the low-coherency series he never had this stability.

The results showed that when stimuli were unpredictable the widely separated stimulus sources induced poorer performance than adjacent, closely spaced ones, which is consistent with other research on the effects of spatial separation of sources cited in the previous section of this chapter on The Observing Response. There were no regularities on which to base anticipation and the subject had no choice but to use an active scanning of head and eyes to pick up the stimulus requirements for response to each source. But, when prediction of events was possible no difference was found between narrow and wide separation of sources. Through anticipation the subject could sidestep the high demands placed on observing by the wide separation. With cues for the center light learned and internal, the subject no longer needed as much visual scanning. He had only to orient to the source where the uncertain event was located – the other could be anticipated. Evidence for anticipation was found in the performance records as response times faster than optimum classical reaction time, as in Poulton's work and our other studies on anticipation. The anticipation effects were most prominent for the fast stimulus rate, which conforms to the study on proprioception and anticipation (Adams & Creamer, 1962b) where discrete stimuli of relatively long duration had poorer anticipation than those of short duration. Longer stimulus duration means longer time between response changes and, in terms of our proprioceptive hypothesis, the proprioceptive trace decays more and is less effective in cuing the motor response.

### *Vigilance*

Students of anticipation should not overlook the domain of vigilance as a source of useful information (Frankmann & Adams, 1962). Vigilance is the study of alertness for the detection of critical signals when they occur. Periods of watchkeeping customarily last one or more hours. With one stimulus source to monitor there can be only temporal uncertainty about when the next signal will occur, but with a visual display having multiple stimulus sources there can also be spatial uncertainty about where a signal will occur. The topic of temporal anticipation runs under the banner of expectancy in the vigilance area (Baker, 1959; Boulter & Adams, 1963; Deese, 1955; Frankmann & Adams, 1962), and considerable work has been done on it. Little has been done on spatial anticipation, however.

Adams and Boulter (1964) have recently reported on man's capability for anticipating where and when a signal will occur in a vigilance task. Their investigation is an example of how anticipation can influence stimulus reception from a complex display. The task had three stimulus sources arranged over 144 degrees and time of pressing a button in response to occurrence of a signal was the measure of performance. The stimulus was the arbitrary number 20 that came on for 5 seconds. If one is so inclined, this task can be viewed as a discrete visual-motor task where motor response time is recorded to stimuli that are spatially dispersed and occur infrequently. Our experimental variables were spatial and temporal uncertainty. There were four groups with all combinations of temporal and spatial certainty and uncertainty. The temporal uncertainty condition (TU) had nine intersignal intervals ranging from 15 to 438 seconds randomized in each 30-minute trial period, but the signals always occurred sequentially on the display from left source to right source and then right to left, etc., so that there was spatial certainty and a subject always knew where a signal would occur. A certainty condition (C) also had spatial certainty but in addition had temporal certainty where the intersignal interval was constant at 195 seconds, which was the mean of the nine intervals for temporal uncertainty. Condition SU had spatial uncertainty with temporal certainty, and condition U had both spatial and temporal uncertainty. A monitoring session was 2.5 hours divided into five uninterrupted 30-minute trials, and a practice session was given in the days immediately before the criterion session to insure learning of the conditions of spatial and temporal uncertainty that would prevail for a subject. Instructions also emphasized the certainties and uncertainties of the task for the subjects of a group.

The findings are presented in Fig. 2. The solid line for each group represents means for all three stimulus sources combined, and the dotted lines are for each stimulus source separately. Performance was directly related to the predictability of events. Group C had events perfectly predictable in time and place, and their performance was best. Group TU had events that were only predictable in place, and they rank next in proficiency.

Uncertainty about where an event will occur is more damaging for performance than uncertainty about when an event will occur because Group SU was distinctly poorer than Group TU. And Group U, which had signals that were both uncertain in time and place, was the poorest of all.

It should be made clear that in this experiment the advantage of temporal certainty (with accompanying event certainty) is somewhat different than in the experiments of Poulton (1952a) and Adams and Creamer (1962a, b). In the latter experiments the subjects anticipated in the fullest sense of the term because they came to jump the gun and key the motor response a bit ahead of stimulus to give response times shorter than the ideal reaction time. In this vigilance study, by virtue of the task and instructions, the subjects always waited until after the event occurred before responding, so here the role of expectancy was probably a heightening of preparatory muscular set. How-

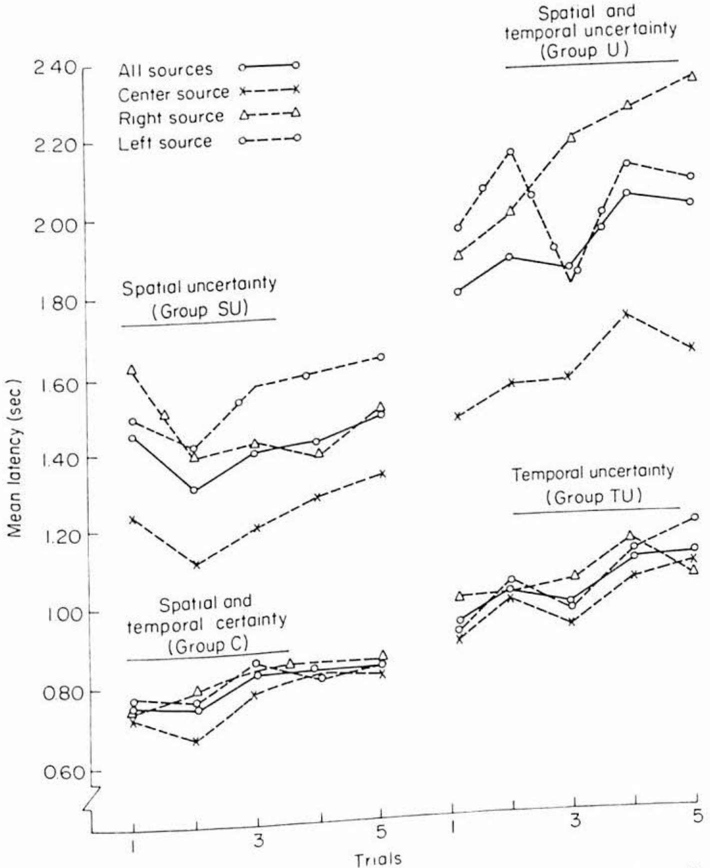


FIG. 2. Performance in a three-source vigilance task as a function of spatial and temporal uncertainty (Adams & Boulter, 1964). See text for fuller explanation.

ever, there is no reason to believe that genuine anticipatory expectancy would have been possible with an interstimulus interval of 195 seconds.

Spatial uncertainty operated in behalf of the observing response in a very positive fashion. With stimulus sources arrayed over 144 degrees there was heavy demand on head and eye movements to scan back and forth, and there was a good chance that a subject would be oriented to the wrong source when a signal occurred. As a result his response latency would be increased by the time it took the observing response to swing around to the source with the signal and detect it. Such visual doubts are absent with spatial certainty because the subject knows where the next event will occur. The only problem for him is when the event will occur, and this problem lessens as temporal uncertainty decreases.

## Limitations in Human Data Processing

Interest in the limits of human data processing has its taproot in engineering psychology where quantification is hoped for the rate at which the human can receive, digest, and respond to stimulus events, just as we quantify the capacities of a digital computer. By knowing these properties of the human being it is hoped he can be used in an informed manner when he is combined with machines in a man-machine system that must conform to specifications for rate of data processing. I have come to regard this topic like guerrilla warfare where the enemy always melts into the bush as you approach him. Considering all the possible stimuli and the ways they can be presented to a subject, all the types of responses that can be required to stimuli, all the mechanisms (e.g., learning, motivation) that can intervene between stimuli and responses, and the great variability that individual differences force upon the data, I see no prospect of locating a tidy set of values that state human capacity and will be useful for the engineering world. The problem is as big as psychology itself. Engineering psychology wants to know how well and how fast a subject responds in a given situation but this is a statement of what all psychologists want to know. Achievement of this goal comes with knowing the laws of psychology, and so the domain of engineering psychology over the long run becomes coexistent with that of psychology at large (Adams, 1961).

Nevertheless, the general line of thinking and research on human data processing has stimulated topics which might otherwise have been bypassed. Like many ideas, hypotheses, and theories that highlight on the scientific panorama and then dim, we are left a bit richer by a residue of new data and ways of viewing behavior. For me, an important residual has come from studies of psychological refractory period or, as it is often called, the hypothesis of the human operator as a one-channel information system. I find the topic of interest for attention and how a subject selects from the total environment that stimulus subset which is given the response of the moment. So far in this chapter a main concern has been with response to multiple stimulus sources that are spatially spread out, so a subject's responding is sequential and one response is usually finished before the next begins. The subject receives stimuli from a source, responds to them, reorients visually to another source, responds to it, etc. But what if the visual observing response is eliminated? How well will the subject respond to two stimulus streams that impinge together or nearly together? Will there be performance degradation, or can a person truly do two things at once? Can he divide his attention, as Woodworth phrased it (1938, p. 708)? Stating the issue colloquially and asking whether a subject can do two things at once fails to capture the germ of the matter, however. We all do two or more things at once, like walking and talking, but can we do them together *without impairment?* Therein lies the issue. If a subject cannot, we can justifiably call him a one-channel system. Expressing the topic in terms of attention, we can ask whether the human subject has internal constraints that allow him to respond fully and without impairment to only one stimulus subset at a time. However one chooses to label the topic, it all becomes another way of asking about mechanisms of selecting a stimulus subset from a stimulus complex for response.

The first relevant empirical evidence was generated by Telford (1931), who found that reaction time to the second of two auditory stimuli was lengthened when the interstimulus interval was decreased to 500 milliseconds. Reasoning through a chain of observations by Thorndike (1927) and Dodge (1927), Telford concluded that after stimulation and response a subject becomes refractory and relatively insensitive to immediate reactivation in a manner analogous to the refractoriness of individual nerve fibers. Craik (1948) apparently was the first to make something conceptually out of all this. He said that there may be a central computing process that acts upon a stimulus when it is received and which controls the discharge down motor nerves. A new stimulus received while this process is going on will be impaired in activating its response.

The common experimental paradigm for testing the one-channel hypothesis has been to present a series of stimulus pairs where the members of the pairs are separated by assigned intersignal intervals. If response to a second member of a pair shows delay when the intersignal interval becomes brief, as compared with a control condition that has response to the second stimulus alone without an immediately prior stimulus to occupy the "computing center," then evidence is mustered for the single channel view.

While all studies have adhered to this paradigm as I have broadly stated it, there has been variation among the details from experiment to experiment and peripheral contamination has been a possibility with some of them. Some studies have used two visual stimuli which, even though seated close together on the display, probably elicit some small action of eye movements which adds to delay in the second response by the amount of time it takes the eyes to shift from one stimulus to the other. Other experiments have used successive responses by the same hand which can add to delay in the second response because of inertia in the arm-hand that would enforce extra time to reactivate this effector. My preference is for the task design used by Davis (1957, 1959) where the two stimuli impinge at different sense modalities to eliminate the visual observing response, and the motor response to each stimulus is by a separate hand. Whatever delay effects now occur can more reasonably be ascribed to central processes. The contribution of peripheral effects in less than optimal task designs is probably not large but if interest is shown in quantification of central delay phenomena, as it has been by our British colleagues, then the quest for a "pure" task design is worthwhile.

It would be helpful to set the stage by citing the findings of a representative experiment by Davis (1957). Using two subjects, Davis obtained a large number of observations from each in a two-dimensional reaction time task where a visual stimulus was presented for key-pressing response by one hand and an audio stimulus for the same response by the other hand. The interstimulus interval varied randomly from 50 to 500 milliseconds, and the interest was whether response time to the second stimulus (audio) would be increased as the intersignal interval decreased. Assessment of increase in reaction time was by comparison with a control series where the task was simple reaction time to the second stimulus alone. A decrement in the second response when interstimulus intervals were short was what Davis found, as have a number of other investigators. For both subjects there was no effect on response to the second signal when intersignal interval was greater than 250 milliseconds. Below 250 milliseconds the decrement steadily increased until at 50 milliseconds reaction time to the second signal was almost twice that of control responses. The human being does indeed seem intolerant of S-R sequences spaced too closely, and a one-at-a-time channel is not an unreasonable hypothesis. More extensive reviews of experiments of this sort are given by Broadbent (1958), Adams (1961, 1964), Welford (1960), and Reynolds (1964).

Hick (1948), Poulton (1950), and Elithorn and Lawrence (1955) have discussed an alternative explanation of data such as these in terms of the expectancy hypothesis. If I divine sentiment in the British literature properly these prophets have been without honor in their own land, but I have found their ideas provocative for my own research and capable of accounting for a good bit of the phenomena. In simplest terms, the expectancy hypothesis says that response delay at short interstimulus intervals is because the subject is not ready, or "set." Over the relatively long stimulus series the subject comes to learn about properties of the random distribution of interstimulus intervals and is thought to be most prepared for responding when the interval is of intermediate length. When a very short interval occurs the subject is not ready and his response to the second stimulus is lengthened. Here is temporal anticipation again, and this usage of expectancy is about the same as in the vigilance area. Vigilance workers have been particularly concerned about expectancy when a distribution of random interstimulus intervals occurs over a long series of stimuli (e.g. Frankmann & Adams, 1962).

An implication of the expectancy hypothesis is that a subject can be a multichannel system and respond to two closely spaced or simultaneous stimuli provided he can acquire appropriate expectancy states. This hypothesis was tested in an experiment (Adams, 1962) by manipulating the random distribution function for interstimulus intervals in the hope of making subjects more expectant for small intervals and show less or no decrement in the second response. The one-channel hypothesis would hold that decrement is a function of interval length defined by the computing center's processing time, not the statistical structure of intervals over the series, so this procedure should discriminate these opposing interpretations. The task presented audio-visual pairs in a long, continuous series, and responding was with two hands. The task had event certainty and the subject always knew what was going to happen at each source, but there was temporal uncertainty about when the visual event would occur and when the audio event would occur after it, and it is the latter uncertainty which was the experimental variable. The audio stimulus lagged the visual one by either 0, 100, 200, 400, or 800 milliseconds for each pair, and the experimental variable was probability of occurrence for each of these values. One group (HU) had high uncertainty of occurrence for intervals, with each interval having equal probability of .20. Rectangular distributions of this sort are generally used in experiments on this topic. A second group (MU) had medium uncertainty, with the probability for the 100-millisecond interval increased to .60 and all other intervals at .10. The third group (LU) had low uncertainty because the probability was .80 for the 100-millisecond interval and .05 for the others. Each subject had unisensory control series and served as his own control. An operations recorder was used to record all stimuli and all responses at the final stage of training, where the analysis was concentrated. If the expectancy hypothesis is correct, our biasing toward the 100-millisecond interval should make subjects more expectant for short intervals and there should be reduced or even no impairment. But, if the one-channel hypothesis holds, then decrement should be found for intervals of 200 milliseconds and less because the computing process for the first S-R pair is considered to operate in this time range.

The results supported the expectancy hypothesis. For intervals of 0, 100, and 200 milliseconds, where impairment is most prominent in other studies, group HU had significant decrement in each case. Group MU had a significant decrement at only the zero interval, and group LU had no significant decrement whatsoever. We made a further analysis of how the distribution of intervals affected the patterning of the two responses to each stimulus pair. We obtained a Difference Response Time measure from our records, which is the time by which response to the second stimulus lagged response to the first stimulus. The means of these measures for the three groups are plotted in Fig. 3 as a function of interstimulus interval. Notice that group HU lags the second response more, presumably because it has a greater frequency of longer intervals and responds with the general expectation that the interval will be rather long. The result for Group HU is longer response times for the second response when interstimulus intervals are small and the decremental effect that is traditionally cited as evidence for the one-channel hypothesis. This trend is essentially absent for the other two groups with their higher frequency of short intervals.

I found this experiment convincing evidence against the one-channel view and for the expectancy position, but an experiment by Adams and Cham-

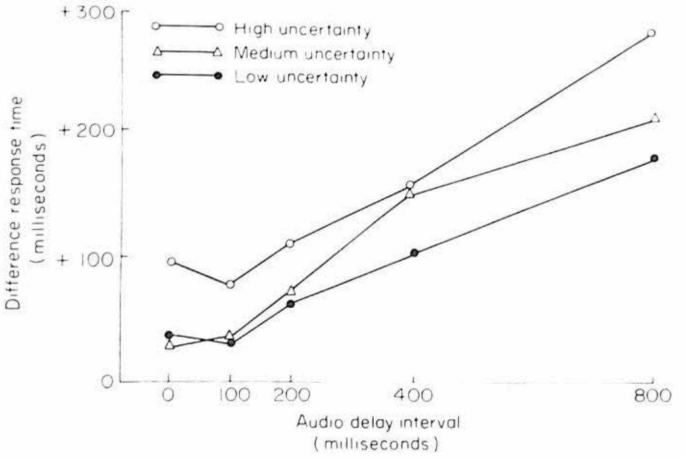


FIG. 3. Difference time between successive responses to two simultaneous or closely spaced stimuli (from Adams, 1962). See text for fuller explanation.

bers (1962) gave data that can be taken as support for the one-channel view provided the definition is restricted. Adams and Chambers used the same task as in the study by Adams (1962) just discussed, and they had temporal certainty throughout, but both event certainty and uncertainty for the audio-visual stimulus pairs. Adams (1962) had event certainty but temporal uncertainty. With certain events a subject always knew exactly what stimulus was going to occur in each modality, and with event uncertainty there were two possible stimuli that could occur in each modality each time. For our purposes here it will suffice to say about the findings that decrement was found when two simultaneous stimuli had event uncertainty, and the possibility emerged that a one-channel view has credence providing the channel is taken to be concerned with the resolution of choices and is considered a decision mechanism, not merely a center that is occupied with stimuli and responses per se. Creamer (1963) took this modified hypothesis as his point of departure and, using essentially the same task as Adams (1962) and Adams and Chambers (1962), had a two-choice situation in each modality for each bisensory pair. His expectation was to find the function relating decrement in the second response as a function of interstimulus interval. Interstimulus intervals were 0, 100, 200, 400, and 800 milliseconds. Events always occurred with complete time certainty so a subject could optimally learn when events would happen. In this way expectancy was ruled out, or at least minimized. All subjects had ideal conditions for acquiring optimal temporal expectancies, so any decrements that were observed could be attributable to the processing of choices and to the smallness of the interstimulus interval. Creamer's results revealed that decrement occurred at the 200-millisecond interval and steadily increased as the interval grew smaller. At the zero interval the decrement was about 50 percent of the baseline reference provided by the unisensory control. In a related study, Adams and Creamer (1962c) found further support for the view that impairment occurs when a subject must process event uncertainties. They found it was necessary to have uncertainty in only one of the channels.

The conclusion from these studies is that decrement in doing two things at once can occur for two main reasons. One reason is temporal uncertainty and the patterning of dual responses as a function of learning about the random distribution of interstimulus intervals over a long series of stimulus pairs. Decrement gotten this way has nothing to do with the one-channel hypothesis and its central computing mechanism because it is a product of how two responses are temporally structured. These findings on temporal learning actually contain the conditions of multi-channel responding because with high event and temporal certainty a subject can indeed do two things at once without impairment (Adams, 1962; Adams & Chambers, 1962). We might surmise that many concurrent activities that we perform in everyday life involve highly stable, predictable sequences with high event and temporal certainty. However, when there is event uncertainty in a channel, and temporal uncertainty has been ruled out as a variable, there is clearcut decrement that is attributable to central restraints in event processing. It is my present conclusion that the human being is a one-channel system provided this channel is defined as a decision center whose function is the resolution of choices. Apparently one reason that a subject attends to one stimulus subset and not another is that he has a very limited capacity for processing choices present in the impinging stimuli. The key to this conclusion has been the unscrambling of event and temporal uncertainty which can produce quite similar behavioral effects for different reasons.

## Conclusions

I hoped to have made it clear at the outset that *attention* need not be elevated to the status of a concept in behavior theory. Rather I tried to emphasize that a minding of the behavioral details that we might carry under the rubric of attention would eventually sharpen the quality of our behavioral laws. Notable among these is closer concern with identification of the stimulus subset to which a response of the moment is given, and I have discussed visual observing, anticipation as an internal cue for the occurrence of a stimulus subset, and how data processing limitations of the human being can determine a stimulus subset that is given unimpaired response. These three aspects of attention are not exhaustive for the many ramifications of these matters but they are aspects which have received their substantive development by investigators of motor behavior and therefore seem particularly worthy of elaboration here. As these issues are attacked with research they will eventually fan out and embrace knowledge in other, more traditional, areas of psychology. Thus, preview of stimuli to occur is clearly a primary research realm for understanding stimulus selection and, in receiving future stimuli, the subject must use short-term memory in holding the stimuli for response. Memory, then, will become a part of the study of attentive behavior, just as other behavioral mechanisms will most assuredly enter our predictive schemata for attention. But, once behavior is explained with concepts like memory, attention falls by the wayside. For a while it will probably be useful to maintain attention as a useful chapter heading for the aggregation of mechanisms that enter the moment-to-moment selection of stimulus subset for response, but I expect that attention may some day die as a useful notion in psychology. Chapter headings are useful for textbooks but not laws and theory. If there is a point of view for this chapter it is that our present approaches to laws and theory are bypassing much of importance that has long been labeled attentive behavior and which has been rather absent from the research scene for a long time. Undoubtedly there are many who still identify attention with old subjective psychology, and who see it outside the pale of modern objective psychology. Such attitudes need recasting. *Attention* may be an unpalatable label for some, and that I understand, but the behavioral phenomena which attention subsumes must stand as a challenge to all psychologists.

## REFERENCES

Adams, J. A. A source of decrement in psychomotor performance. *J. exp. Psychol.*, 1955, **49**, 390–394.

Adams, J. A. Human tracking behavior. *Psychol. Bull.*, 1961, **58**, 55–79.

Adams, J. A. Test of the hypothesis of psychological refractory period. *J. exp. Psychol.*, 1962, **64**, 280–287.

Adams, J. A. Motor skills. *Annu. Rev. Psychol.*, 1964, **15**, 181–202.

Adams, J. A., & Boulter, L. R. An evaluation of the activationist hypothesis of human vigilance. *J. exp. Psychol.*, 1962, **64**, 495–504.

Adams, J. A., & Boulter, L. R. Spatial and temporal uncertainty as determinants of human vigilance. *J. exp. Psychol.*, 1964, **67**, 127–131.

Adams, J. A., & Chambers, R. W. Response to simultaneous stimulation of two sense modalities. *J. exp. Psychol.*, 1962, **63**, 198–206.

Adams, J. A., & Creamer, L. R. Anticipatory timing of continuous and discrete responses. *J. exp. Psychol.*, 1962, **63**, 84–90. (a)

Adams, J. A., & Creamer, L. R. Proprioception variables as determiners of anticipatory timing behavior. *Hum. Factors*, 1962, **4**, 217–222. (b)

Adams, J. A., & Creamer, L. R. Data processing capabilities of the human operator. *J. engng. Psychol.*, 1962, **1**, 150–158. (c)

Adams, J. A., & Webber, C. E. The organization of component response error events in two-dimensional visual tracking. *J. exp. Psychol.*, 1961, **61**, 200–212.

Adams, J. A., & Webber, C. E. Monte Carlo model of tracking behavior. *Hum. Factors*, 1963, **5**, 81–102.

Adams, J. A., & Xhignesse, L. V. Some determinants of two-dimensional tracking behavior. *J. exp. Psychol.*, 1960, **60**, 391–403.

Adams, J. A., Stenson, H. H., & Humes, J. M. Monitoring of complex visual displays, II: effects of visual load and response complexity on human vigilance. *Hum. Factors*, 1961, **3**, 213–221.

Arnoult, M. D. The curvilinearity of normal eye movements. In G. Finch and F. Cameron (Eds.), *Air Force human engineering, personnel, and training research*. Washington, D.C.: Nat. Acad. Sci. – Nat. Res. Council (Publ. No. 516), 1958, pp. 140–146.

Bahrick, H. P. An analysis of stimulus variables influencing the proprioceptive control of movements. *Psychol. Rev.*, 1957, **64**, 324–328.

Bakan, P. Discrimination decrement as a function of time in a prolonged vigil. *J. exp. Psychol.*, 1955, **50**, 387–390.

Baker, C. H. Towards a theory of vigilance. *Canad. J. Psychol.*, 1959, **13**, 35–42.

Baker, C. H. Observing behavior in a vigilance task. *Science*, 1960, **132**, 674–675.

Boring, E. G. *A history of experimental psychology*. New York: Appleton, 1929.

Boulter, L. R., & Adams, J. A. Vigilance decrement, the expectancy hypothesis, and intersignal interval. *Canad. J. Psychol.*, 1963, **17**, 201–209.

Briggs, G. E., & Howell, W. C. On the relative importance of time sharing at central and peripheral levels. *Nav. Train. Devices Center tech. Rep.*, 1959, No. 508-2.

Broadbent, D. E. *Perception and communication*. Oxford: Pergamon Press, 1958.

Craik, K. J. W. Theory of the human operator in control systems, II: man as an element in a control system. *Brit. J. Psychol.*, 1948, **38**, 142–148.

Creamer, L. R. Event uncertainty, psychological refractory period, and human data processing. *J. exp. Psychol.*, 1963, **66**, 187–194.

Davis, R. The human operator as a single channel information system. *Quart. J. exp. Psychol.*, 1957, **9**, 119–129.

Davis, R. The role of 'attention' in the psychological refractory period. *Quart J. exp. Psychol.*, 1959, **11**, 211–220.

Deese, J. Some problems in the theory of vigilance. *Psychol. Rev.*, 1955, **62**, 359–368.

Delafresnaye, J. F., (Ed.) *Brain mechanisms and consciousness*. Springfield, Ill.: Thomas, 1954.

Dodge, R. A note on Professor Thorndike's experiment. *Psychol. Rev.*, 1927, **34**, 237–240.

Elithorn, A., & Lawrence, C. Central inhibition: some refractory observations. *Quart. J. exp. Psychol.*, 1955, **7**, 116–127.

Estes, W. K. The statistical approach to learning theory. In S. Koch (Ed.), *Psychology: a study of a science*. Vol. 2. New York: McGraw-Hill, 1959. pp. 380–491.

Fitts, P. M., & Simon, C. W. Some relations between stimulus patterns and performance in a continuous dual-pursuit task. *J. exp. Psychol.*, 1952, **43**, 428–436.

Ford, A. Foundations of bioelectronics for human engineering. *USN Electronics Lab. res. Rep.*, 1957, No. 761.

Ford, A., White, C. T., & Lichtenstein, M. Analysis of eye movements during free search. *J. opt. Soc. Amer.*, 1959, **49**, 287–292.

Frankmann, J. P., & Adams, J. A. Theories of vigilance. *Psychol. Bull.*, 1962, **59**, 257–272.

Gifford, R. N. Tracking performance of the human operator with advanced and delayed visual displays. Unpublished master's thesis, Univer. of California (Los Angeles), 1963.

Guthrie, E. R. Association by contiguity. In S. Koch (Ed.), *Psychology: a study of a science*. Vol. 2. New York: McGraw-Hill, 1959, pp. 158–195.

Hebb, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.

Hick, W. E. The discontinuous functioning of the human operator in pursuit tasks. *Quart. J. exp. Psychol.*, 1948, **1**, 36–51.

Hull, C. L. *Principles of behavior*. New York: Appleton, 1943.

Irion, A. L. The relation of 'set' to retention. *Psychol. Rev.*, 1948, **55**, 336–341.

Jasper, H. H., (Ed.) *Recticular formation of the brain*. Boston, Mass.: Little, Brown, 1958.

Jones, R. E., Milton, J. L., & Fitts, P. M. Eye fixations of aircraft pilots, I: a review of prior eye-movement studies and a description of a technique for recording the frequency, duration, and sequences of eye fixations during instrument flight. *USAF Air Materiel Command tech. Rep.*, 1949, No. 5837.

King, W. J. Coordination in a complex tracking task as a joint function of spatial separation and predictability of stimuli. Unpublished master's thesis, Univer. of Illinois, 1961.

Leonard, J. A. Advance information in sensori-motor skills. *Quart. J. exp. Psychol.*, 1953, **5**, 141–149.

Lindsley, D. B. Psychophysiology and motivation. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1957*. Lincoln, Univer. of Nebraska Press, 1957, pp. 44–105.

McFarland, R. A., Holway, A. N., & Hurvich, L. M. Studies of visual fatigue. Grad. Sch. Bus. Admin. Harvard Univer., 1942.

Mackworth, J. F., & Mackworth, N. H. Eye fixations recorded on changing visual scenes by the television eye-marker. *J. opt. Soc. Amer.*, 1958, **48**, 439–445.

Mackworth, N. H. Researches on the measurement of human performance. In H. W. Sinaiko (Ed.), *Selected papers on human factors in the design and use of control systems*. New York: Dover Publications, 1961. pp. 174–231.

Magoun, H. W. *The waking brain* (2nd ed.). Springfield, Ill.: Thomas, 1963.

Malmo, R. B. Activation: a neuropsychological dimension. *Psychol. Rev.*, 1959, **66**, 367–386.

Moruzzi, G., & Magoun, H. W. Brain stem reticular formation and activation of the EEG. *EEG clin. Neurophysiol.*, 1949, **1**, 455–473.

Oswald, I. *Sleeping and waking*. Amsterdam: Elsevier, 1962.

Paschal, F. C. The trend in theories of attention. *Psychol. Rev.*, 1941, **48**, 367–382.

Poulton, E. C. Perceptual anticipation and reaction time. *Quart. J. exp. Psychol.*, 1950, **2**, 99–112.

Poulton, E. C. The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, 1952, **43**, 295–302. (a)

Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229 (b)

Poulton, E. C. Learning the statistical properties of the input in pursuit tracking. *J. exp. Psychol.*, 1957, **54**, 28–32. (a)

Poulton, E. C. On prediction in skilled movements. *Psychol. Bull.*, 1957, **54**, 467–478. (b)

Poulton, E. C. On the stimulus and response in pursuit tracking. *J. exp. Psychol.*, 1957, **53**, 189–194. (c)

Reynolds, D. Effects of double stimulation: temporary inhibition of response. *Psychol. Bull.*, 1964, **62**, 333–347.

Rossi, G. F., & Zanchetti, F. The brain stem reticular formation. *Arch. ital. Biol.*, 1957, **95**, 199–435.

Samuels, I. Reticular mechanisms and behavior. *Psychol. Bull.*, 1959, **56**, 1–25.

Sanders, A. F. *The selective process in the functional visual field*. Soesterberg, The Netherlands: Report of the Institute for Perception R.V.O.–T.N.O., 1963.

Senders, V. L. The effect of number of dials on qualitative reading of a multiple dial panel. *USAF WADC tech. Rep., 1952, No. 52-39.*

Sharpless, S., & Jasper, H. H. Habituation of the arousal reaction. *Brain*, 1956, **79**, 655–680.

Telford, C. W. Refractory phase of voluntary and associative responses. *J. exp. Psychol.*, 1931, **14**, 1–35.

Thorndike, E. L. The refractory period of associative processes. *Psychol. Rev.*, 1927, **34**, 234–236.

Titchener, E. B. *Lectures on the elementary psychology of feeling and attention.* New York: Macmillan, 1908.

Underwood, B. J. Stimulus selection in verbal learning. In C. N. Cofer and B. S. Musgrave (Eds.), *Verbal behavior and learning.* New York: McGraw-Hill, 1963. pp. 33–48.

Watson, J. B. Psychology as the behaviorist views it. *Psychol. Rev.*, 1913, **20**, 158–177.

Waugh, N. C. Free versus serial recall. *J. exp. Psychol.*, 1961, **62**, 496–502.

Welford, A. T. The measurement of sensory-motor performance: survey and reappraisal of twelve years' progress. *Ergonomics*, 1960, **3**, 189–230.

White, C. T., & Ford, A. Eye movements during simulated radar search. *J. opt. Soc. Amer.*, 1960, **50**, 909–913.

Woodrow, H. Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. pp. 1224–1236.

Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.

Wyckoff, L. B., Jr. The role of observing responses in discrimination learning: Part I. *Psychol. Rev.*, 1952, **59**, 431–442.

Zuercher, J. D. The effects of extraneous stimulation on vigilance. *Hum. Factors*, 1965, **7**, (in press).

# Attention and Vigilance as Mechanisms of Response

## Comments on Professor Adams' Paper

EARL A. ALLUISI
*University of Louisville*

The work of science is largely the search for mechanisms. If the meteorologist in his search for an understanding of the weather seeks to define the mechanisms responsible for storms, he is acting in the prime role of a scientist. If he concludes that atmospheric pressure change is the mechanism principally responsible for high winds, there is scientific understanding of his meaning and scientific approval of his approach.

Now, if behavior (and more specifically, animal and human behavior) is the fundamental domain of psychology, it is understandable that psychologists who want to advance the scientific study of behavior seek mechanisms. They seek the mechanisms of behavior, or the mechanisms of response, since we know behavior only through response. In this psychology is following its parent sciences, most notably physics.

On the other hand, the inclusion of this topic here might be viewed as a demonstration of great egotism or of great naïveté. What physicist would dare face a gathering of colleagues to discuss the "mechanisms of matter and energy"? He might agree to address a group of junior high school students on such a topic, if he were assured that the students had not yet studied high school physics. Yet we psychologists do dare to include "mechanisms of response" among our topics! This is, perhaps, a sign of the relative youth of psychology as a science—an admission that we know little, indeed, of these mechanisms.

In classical scientific psychology, we have approached the study of our mechanisms of behavior through the isolation of stimulus variables, response variables, or organismic variables (cf. Woodworth & Schlosberg, 1954, pp. 1–7). Our stimulus (S) variables have included the modality of stimulation and the source of the stimulating matter or energy—the *physics* of *psychophysics*—as in the wavelength of light and the frequency of sound; the chemical composition of the liquids we taste and of the gases we smell; the intensity and duration of the stimulus, and in certain cases, its area, shape, and spatial relations such as distance, size, direction, etc. When we progress from the static situation to the dynamic, we add new and complicating dimensions related to various kinds of movement, change, and rate of change; these interact with our initial dimensions and we begin to look at changes in frequency or wavelength, rates of change in intensity or in duration, etc.

Our response (R) variables have traditionally emphasized the criterion or measure of performance employed—speed, accuracy, the difficulty level achieved, the likelihood of occurrence of a response, or the strength of a response. We have attended all too little to the task imposed on the subject and to the relation between the performance of that task and the criterion of performance. For example, Stevens (1957) established a task that is completely different from the classical psychophysical tasks of Fechner (see Woodworth & Schlosberg, 1954, pp. 192–266; Guilford, 1954), yet much has been written about the differences obtained with his methods relative the classical methods and of the scientific validity of his "power law" vs. the Fechnerian logarithmic law (see Stevens, 1957; Luce, 1959). Perhaps we should expect different laws to apply to different behaviors, or, rather, to behaviors that are observed in the performance of different tasks!

The organismic (O) variables of scientific psychology have emphasized the reality of individual differences (some of which have been related to the organism's age, "ability," organic state, etc.), to the level of skill or learning attained, and to other factors believed to affect performance such as drive, incentive, inhibition, fatigue, interest, and various conditions in states of which the organism is hypothesized to be.

We have been troubled, as a science, with the problem of interactions among these "S," "O," and "R" variables. In fact, we have generated whole classes of variables or mechanisms that are dependent logically on elements of two or more of the three. Thus, we have "motivation" as something that should be an organismic variable, but which is clearly affected by the task and by both the stimulus and response variables that help define the task. "Attention" is another one of these "situational variables," as we shall call them here, and thus it is not inappropriate that we consider it a suitable topic in discussing mechanisms of response.

## Attention and Research on Vigilance

As Adams (pp. 169–171) has pointed out, the study of attention has had a long history in experimental psychology. *Attention*, *set*, and *Aufgabe* are terms that once were central themes of psychological research (cf. Boring, 1957). Renewal of interest in this area is generally said to date from the publication of Mackworth's monograph on "Researches on the Measurement of Human Performance" in 1950. This was the paper in which "vigilance tests" were first reported as such. Of course, we can find prior anticipations in the literature, notably Ditchburn (1943) and Mackworth (1948) in England, Best, Solandt, and Partridge (1950) in Canada, and Lindsley (1944) in the United States; but is seems correct to say that it was not until after the publication of the Mackworth (1950) monograph that general psychological research interest was aroused in the topic of vigilance or watchkeeping behavior.

The terms *vigilance, watchkeeping, monitoring,* and *signal detecting* have all been used to describe this area of study. The area is now generally taken to be concerned with "... the attentiveness of the subject and his capability for detecting changes in stimulus events over relatively long periods of sustained observation. . . ." (Frankmann & Adams, 1962, p. 257).

Mackworth used the term "vigilance" in a more specific sense. He used it in a sense that he credits to H. Head, but with psychological rather than physiological emphasis: *vigilance* meant to describe states of psychological readiness to perform adapted and purposive acts (cf. Mackworth, 1950, p. 8). Vigilance was a mechanism of response.

The task in which the subject was tested was a watchkeeping task; the situation to which generalization was desired was the watchkeeping functions of certain military jobs, especially the job of the radar lookout. Vigilance was the mechanism of response believed to be principally responsible for the watchkeeper's performance: his state of readiness to respond correctly to the occurrence of a critical signal. *Monitoring* and *signal detecting* seem to be terms aimed at describing what the subject has to do in a watchkeeping task designed to permit the study of vigilance.

Our usages of these terms have been less consistent than indicated by the preceding paragraphs, and this probably reflects something of a general state of confusion with regard to the concepts involved. Indeed, the looseness with which we have used these terms and confounded these concepts may well have contributed significantly to the confusion. As Frankmann and Adams (1962) have pointed out in their review of our contemporary theories of vigilance, a major shortcoming is the ". . . casualness of formulation that makes the definitive testing of implications rather difficult. . ." (Frankmann & Adams, 1962, p. 268).

We shall try to be consistent in our usage throughout the remainder of this paper. Thus, a *watchkeeping* task is one in which the subject is asked to *monitor* a display in order to *detect signals;* the subject's *vigilance* is believed to be the mechanism principally responsible for his watchkeeping behavior! With these concepts in mind, let us see how research on vigilance is related to classical psychological research on attention.

### THE CLASSICAL STUDY OF ATTENTION

In the classical study of attention (for a good review see Woodworth & Schlosberg, 1954, ch. 4, pp. 72–106), a differentiation is made between *free* and *controlled* attention. In free attention there is no assigned task, whereas in controlled attention, as in controlled association, the subject is given a set. This should probably be treated as a dimensional affair, rather than as an experimental condition admitting of only two levels; it should be regarded as a matter of the degree to which instructions, set, *Aufgabe*, or schemata determine the orienting, observing, and attentive behavior of the subject. In modern terms it can be viewed as the degree to which the goal of the attentive process is uncertain. Instructions are meant to reduce the uncertainty; set reduces the uncertainty; needs and certain drive states probably reduce the uncertainty. The degree to which attention is free or controlled in watchkeeping tasks generally has not been overtly specified. To the extent that it has not been specified and controlled, our research on vigilance would be classified as "sloppy" research by the worker who takes a more classical approach to the study of attention.

Five major areas of research on attention are usually identified: the first is (1) the isolation of the *determiners of attention*. This area includes (1.1) the specification of the "conspicuity" of various warning displays and emergency controls by the engineering psychologist, (1.2) the evaluation of the attention-getting value of various aspects of stimulation such as the use of chromatic vs. achromatic visual stimuli by the experimental psychologist, and (1.3) measurement of the effects (particularly on retention) of stimulus patterns thought to be satisfying to personal needs such as security, success, sex, curiosity, and prestige by the advertising specialist and the consumer-behavior psychologist. Research on vigilance has not followed this general line of study.

The second major area of research on attention is that which studies (2) the *shifting and fluctuation of attention*. The basic method is to record changes in the subject's responses over a period of time while the stimulus field is held constant. The (2.1) "ordinary" shifting of both free and controlled attention has been studied, principally through measurement of the orienting response (eye movements in reading or in the "free" viewing of pictures). Special conditions of attentive shifting also have been studied as in (2.2) the perceived oscillations of ambiguous figures, (2.3) retinal rivalry, (2.4) sensorimotor fluctuations of threshold, adaptation, and fatigue, and (2.5) fluctuations of efficiency during continuous work. Research on vigilance apparently should be related to this last area of research on attention, i.e., to studies of the fluctuations of efficiency during continuous work, and we shall look more closely at this area after we have completed our catalog of the five classical approaches to the study of attention.

The third major area is that of (3) *distraction*. While a subject is engaged in a certain task, making appropriate responses to certain stimuli, irrelevant stimuli are introduced and their distracting effects measured through changes in the speed or accuracy of the subject's task-related responses. The distracting stimuli usually are selected from a different modality than that of the task-related stimuli so as to avoid problems of sensory interference or masking. In general, we have concluded that these irrelevant stimuli may or may not have a measured detrimental effect on performance dependent upon many other variables such as motivation, effort expended, and various specifics of the task and of the distracting stimuli. In addition, work on intersensory facilitation appears to suggest that at least under some conditions, "distracting" stimuli may actually assist performance; the extent to which this might be mediated through the attentive process, or to which it might affect attention as a mechanism of response-mediating behavior, is not clear. It complicates our concept of distraction. Nevertheless, we should note that research on vigilance is but little related, if related at all, to the mainstream of research on distraction.

The fourth major area of classical research on attention is that of (4) *divided attention* or the time sharing of attentive processes. The question of how man should be characterized in his information-handling activities—whether he should be viewed as a single-channel or multiple-channel information-handling system—is a modern restatement of the fundamental question attacked in research on divided attention. It is also related to the question of channel capacity, especially in terms of the rate of information transmission. Research on vigilance is not directly related to this question, though it may be indirectly related as Adams (pp. 190–196) points out in the last section of his paper.

The final major area is that of (5) the span of attention, or as it is more frequently called, the *span of apprehension*. For philosophical reasons this was one of the earliest areas of psychological experimentation; it seemed important to know whether more than one object could be apprehended at a given time. The area has lost some of its importance for psychology in modern times, and we would now characterize it as dealing with the amount of information in absolute judgments of numerosity. To some extent this is also related to the question of channel capacity, but here in terms of the amount of information rather than the rate of information transmission as was the case with distraction. Research in this area seems quite unrelated to research on vigilance.

Thus, it appears that, of the areas of classical study, only one subarea of research on attention has direct relevance to research on vigilance. That is the area of shifting attention as studied through measurements of fluctuations of efficiency during continuous work. Let us now look at the watchkeeping task that is used in the study of vigilance so that we may understand it well enough to make informed judgments as to the operational relation of the two mechanisms of response, vigilance and attention. We need to clarify the extent to which the two concepts are identical and the extent to which they differ.

## THE WATCHKEEPING TASK

Let us picture the general task of watchkeeping much as an historical novelist might. Think of the days of sailing vessels. and think of the lookout's job as demonstrating the elements of the watchkeeping task. The man in the crow's nest had the job of sensing anything atypical. He did not know *when* it would appear, he did not know *where* it would appear, and he did not know *what* would appear (or whether he would sense it by ear or eye, or even whether he would be able to detect it, much less identify it, if, when, and where it did appear).

That is to say, one might characterize the stimulus aspects of the watchkeeping task as involving *temporal* uncertainty, *spatial* uncertainty, and what (for the lack of a better name) we shall call *stimulus* uncertainty. The task of the watchkeeper is to monitor the display (be it the world about him, as in the case of the lookout in the crow's nest of the sailing vessel, or a specific electronic display, as in the case of the operator of a radar detection system) for some period of time or length of watch, and to report the occurrence of certain signals. The report may be asked for in simple signal-detection terms, or various degrees of identification may be required prior to an overt response; in either case, the signal must be detected, so we can say that all watchkeeping tasks require signal detections. The criteria of performance that may be used are the proportion of correct detections or its complement, the proportion of missed signals, reaction time to correct detections, and the number of false responses.

That the kind and degree of uncertainty of the stimulus—the critical signal to which a response is desired—affect the performance of the watchkeeper, perhaps even define the difficulty of the watchkeeping task, seems logical. This has been recognized since the earliest of watchkeeping studies, and yet our manipulation of these aspects of the task has been something less than systematic.

For example, in his 1950 monograph, Mackworth reported that the deterioration in correct signal detections which typically occurred in a synthetic radar task could be delayed for an hour by increasing the brightness of the critical signals relative to their background (Mackworth, 1950, pp. 53–63). This implies that watchkeeping performance is a decreasing function of stimulus uncertainty; i.e., as the signal-to-noise ratio of the critical signal in a watchkeeping task is increased, performance efficiency can also be expected to increase. This generalization is consistent with the findings of Adams (1956), Bartlett, Beinert, and Graham (1955), and Jenkins (1953).

The effects of temporal uncertainty have been much more widely investigated, though apparently with little more systematization. Lindsley (1944) used a signal rate that was 10 times faster than the 24-signals-per-hour rate used subsequently by Mackworth (1950); Lindsley found no decrements in watchkeeping performance until the third day of work, whereas Mackworth found them in the second half-hour of the first watch. Also, both Ditchburn (1943) and Mackworth (1950) reported that the proportions of missed signals changed when intersignal intervals were changed, and later studies made it apparent that increases either in signal frequency or in signal regularity affect performance efficiency in the watchkeeping task (e.g., Baker, 1959a, b, c; Bartlett, *et al.*, 1955; Broadbent, 1958; Deese, 1955; Deese & Ormond, 1953; Jenkins, 1958). Smith (1961) has proposed a psychophysical approach and, using three levels of signal frequency (6, 24, and 96 signals per hour) and three levels of variability (.01, .10, and 1.00 as coefficients of variability of intersignal intervals), he reports a coefficient of correlation of +.959 between reaction time to correct detections and a derived informational measure of the temporal uncertainty of the signal—a measure that combines both frequency and variability of intersignal intervals. It appears safe to conclude that watchkeeping performance is a decreasing function of the temporal uncertainty of the signal; i.e., performance efficiency decreases as the signals become less regularly spaced in time and as they become farther spaced in time.

It was again at least as early as the 1950 Mackworth monograph that spatial uncertainty was found to affect performance in the watchkeeping task. With the synthetic radar task, Mackworth (1950, especially pp. 58–59) found a progressive fall in the incidence of missed signals with rising spatial expectancy—i.e., with decreased spatial uncertainty. Nicely and Miller (1957) scheduled two thirds of the signals in a simulated radar task to appear in one quadrant with the remaining one third distributed equally over the other three quadrants; a relatively high rate of 108 signals per hour was used. They found that proportionally more signals were detected in the high-density quadrant than in the low-density quadrants; they found also that decrements in performance appeared after the first 15 minutes of the watch for the low-density areas, but that no decrements occurred for the high-density quadrants. The results of Adams and Boulter (1964) also lead to the conclusion that watchkeeping performance is a decreasing function of the spatial uncertainty of the signal, and this conclusion is further suggested by the prior results of Bartlett *et al.* (1955) and of Deese and Ormond (1953).

As an aside we would like to comment that with all this evidence to indicate the importance of spatial, temporal, and stimulus uncertainty to watchkeeping performance, it is quite surprising to find so few studies in which functional relations have been sought. In fact, the only study that has attempted to measure the functional relation between performance and one of these three aspects of the watchkeeping task is that of Smith (1961), in which reaction time was related to the temporal uncertainty of the critical signal.

Thus, we may say that the signal to be detected in a watchkeeping task may vary in temporal uncertainty, spatial uncertainty, or stimulus uncertainty. Further, we may conclude that watchkeeping performance will vary as a decreasing function of increases in any of these three sources of uncertainty. Since vigilance is conceived as the mechanism principally responsible for the watchkeeper's performance, it follows that vigilance, too, will vary as a function of spatial, temporal, and stimulus uncertainty.

## VIGILANCE AND ATTENTION

Although Mackworth (1950, pp. 7–11), in the introduction of his monograph on watchkeeping, associated vigilance with attention (remember, *vigilance* for Mackworth refers to the watchkeeper's psychological readiness to respond), it is Broadbent (1953, 1958) who really champions the association. Because the association seems logical, other researchers have accepted it, apparently without comparing the operations that define the two mechanisms of response.

In the typical study of vigilance, the critical signal to be detected may have temporal, spatial, and stimulus uncertainty; usually, spatial and stimulus uncertainty are low (the subject knows where to look and what to look for), whereas temporal uncertainty is high (he does not know when the signal will appear, and the signal appears only infrequently at irregular intervals). In the typical study of shifting attention where fluctuations of efficiency during continuous work are measured, these variables have not been involved (cf. Woodworth & Schlosberg, 1954, pp. 81–84).

In terms of the operations employed in the two situations, then, we find no reason to conclude that the two mechanisms of response "vigilance" and "attention" should be equated. If we mean to broaden the concept of attention by having it now include vigilance as a subconcept, we may do so; but before we agree to this, might it not be well for us to ask what might be gained in so doing? If we accept vigilance as the mechanism of response principally responsible for watchkeeping performance, what do we gain by saying further that it is somehow related generally to attention, though not directly related to any of the attentive processes that have been identified as mechanisms of other responses? The association of vigilance and attention, at least as mechanisms of response, is a topic that deserves further thought and discussion.

## Vigilance and Arousal Theory

Adams (pp. 171–175) points out that arousal theory or the "activation hypothesis" has been advanced as an explanation of watchkeeping performance, and he makes it clear (cf. p. 172) that he believes the association to be an unhappy one. We agree with him in this opinion, and we would like to present additional evidence in support of this position.

The temporal course of watchkeeping performance typically demonstrates a decline in efficiency as time on duty increases (e.g., see Mackworth, 1950). The "activation hypothesis" has been advanced as a possible explanation of this decline (cf. Broadbent, 1958; Frankmann & Adams, 1962). This application of arousal theory (Hebb, 1955; Scott, 1957) is based on the idea that a temporally correlated loss in performance efficiency will occur in the presence of a uniform sensory environment, and that the typical watchkeeping task places the subject in such an environment.

The converse should hold, of course. That is to say, the presence of extraneous stimuli (up to some optimum level, at least) should serve to maintain or restore efficient watchkeeping behavior to the extent that these extraneous stimuli provide a varied sensory environment. However, in a direct test of this—a test in which watchkeeping performance on an auditory vigilance task was measured over the period of a 4-hour watch spent in a varied sensory environment that required high multiple-task activity of the subject—the typical watchkeeping decrements still occurred (Alluisi & Hall, 1963). A later study indicated further that performance efficiency in watchkeeping varied *inversely* with variations in sensory input (Hawkes, Meighan, & Alluisi, 1964).

In fact, Haider, Spong, and Lindsley (1964) have recently reported the only direct evidence supporting the contention that arousal is even remotely related to vigilance within the limits of the typical watchkeeping study, and their evidence is difficult to interpret. They recorded computer-averaged cortical potentials evoked by nonsignal and signal stimuli in a visual watchkeeping task (a high signal rate of 120 signals per hour was used); they found that as detection efficiency decreased with time, the amplitude of evoked responses to *nonsignal* stimuli also decreased and latency increased. The evoked potentials to the *signal* stimuli differed according to whether the signal was detected or missed, but neither of these potentials (nor their difference) changed systematically over the period of the watch. In short, the data indicated changes attributable to nonsignal stimuli, but no change attributable to signal stimuli.

In many cases, performance has been found to be significantly affected by various conditions, while the physiological indicants of arousal fail to show any change (cf. Malmo, 1963, pp. 2–6). Other studies have found changes in the physiological indicants, without changes in performance (e.g., Alluisi, Chiles, & Hall, 1964a; Alluisi, Chiles, Hall, & Hawkes, 1963). These findings will clearly limit generalizations concerning the relations between physiological activation and performance, as Elliott (1965) has recently pointed out.

This is more than the difficulty of the insensitivity of physiological measures as indices of performance efficiency and performance reserves (cf. Alluisi, Chiles, & Smith, 1964b; Alluisi, 1964), and the arousal theorists seem to be leading the way in revising the activation concept (cf. Malmo, 1958, 1959, 1962, 1963). It is beginning to appear as though arousal theory may have relevance to *autonomic-linked behavior* (such as variations in motivation, in level of induced tension, and in degree of wakefulness), but not to *cortical mechanisms;* Malmo says as much (1963, p. 5).

## Mechanics of Mechanisms of Response

Still, a more fundamental question may be raised; namely, to what extent should we try to "explain" how a mechanism of response works? What level of explanation is needed for the concept of the mechanism to advance our scientific knowledge? Is it sufficient to say that vigilance is the mechanism principally responsible for watchkeeping behavior, and then to attempt to write the functional relations between behavior and the task and situational variables—the mechanics or relations that operationally define the mechanism, vigilance—or must we have more? Is our scientific knowledge advanced by the postulation of mechanisms that are supposedly responsible for the manner of operation of the vigilance mechanism, which, in turn, is postulated as responsible for the watchkeeping behavior that we observe at first hand? If so, how far do we go, and when do we stop?

We have proposed that watchkeeping performance varies as a function of the subject's vigilance. We have also proposed that the subject's vigilance will vary as a function (1) of reinforcement and extinction (Mackworth, 1950), (2) of subjective expectancies (Baker, 1959a; Deese, 1955), (3) of the channel capacity of man and his filtering characteristics (Broadbent, 1958), (4) of operant extinction (Holland, 1958), and (5) of the subject's level of physiological arousal (Deese, 1955; Frankmann & Adams, 1962; McGrath, 1963). It would appear that our second-order mechanisms are premature, and will remain so until we have defined the first-order vigilance mechanism.

And how are we to define our first-order mechanism except by making clear its mechanics – by drawing the functional relations? How, other than by showing that vigilance as measured by such-and-so criterion is a this-kind-of function of that-sort-of variation in the watchkeeping task? Nor is what we are proposing to be classed simply as either an antitheoretical or an atheoretical position; rather, it is a call for scientifically meaningful, quantitative, functional theories not merely of vigilance as the mechanism of watchkeeping behavior, but also of all other psychological mechanisms of behavior. We hold that anything less is not science; it may be art, it may be humanistic, it may be philosophy, it may even be theology, but it is not science.

## REFERENCES

Adams, J. A. Vigilance in the detection of low intensity visual stimuli. *J. exp. Psychol.*, 1956, **52**, 204–208.

Adams, J. A., & Boulter, L. R. Spatial and temporal uncertainty as determinants of human vigilance. *J. exp. Psychol.*, 1964, **67**, 127–131.

Alluisi, E. A. Comments on the measurement of performance decrement. In *Dept. Army Rep. Tenth Annu. Army Hum. Factors Res. Develpm. Conf.* U.S. Army, Office of the Chief of Res. Develpm., Washington, D.C., 1964, Pp. 213–226.

Alluisi, E. A., & Hall, T. J. Declines in auditory vigilance during periods of high multiple-task activity. *Percept. mot. Skills*, 1963, **16**, 739–740.

Alluisi, E. A., Chiles, W. D., Hall, T. J., & Hawkes, G. R. Human group performance during confinement. *USAF Aerospace Med. Res. Lab. Tech. Doc. Rep.* No. AMRL-63–87, Wright-Patterson AFB, Ohio, 1963.

Alluisi, E. A., Chiles, W. D., & Hall, T. J. Combined effects of sleep loss and demanding work-rest schedules on crew performance. *USAF Aerospace Med. Res. Lab. tech. Doc. Rep.*, No. 64–63, Wright-Patterson AFB, Ohio. (a)

Alluisi, E. A., Chiles, W. D., & Smith, R. P. Human performance in military systems: some situational factors influencing individual performance. *Amer. Psychologist*, 1964, **19**, 537 (Abstract). (b)

Baker, C. H. Attention to visual displays during a vigilance task: II. Maintaining the level of vigilance. *Brit. J. Psychol.*, 1959, **50**, 30–36. (a)

Baker, C. H. Three minor studies of vigilance. *Def. Res. Med. Lab. (Canada) Rep.*, No. 234-2. Toronto, Ont., 1959. (b)

Baker, C. H. Towards a theory of vigilance. *Canad. J. Psychol.*, 1959, **13**, 35–42. (c)

Bartlett, S. C., Beinert, R. L., & Graham, J. R. Studies of vigilance fatigue and efficiency in radar observation. *USAF RADC tech. Rep.* No. 55-100, 1955.

Best, C. H., Solandt, O. M., & Partridge, J. Wartime reports from the University of Toronto. Unpublished report cited by Mackworth (1950).

Boring, E. G. *A history of experimental psychology.* New York: Appleton, 1957.

Broadbent, D. E. Classical conditioning and human watchkeeping. *Psychol. Rev.*, 1953, **60**, 331–339.

Broadbent, D. E. *Perception and communication.* New York: Macmillan (Pergamon), 1958.

Deese, J. Some problems in the theory of vigilance. *Psychol. Rev.*, 1955, **62**, 359–368.

Deese, J., & Ormond, E. Studies of detectability during continuous visual search. *USAF WADC tech. Rep.* No. TR-53-86, 1953.

Ditchburn, R. W. Some factors affecting efficiency of work of lookouts. *Admiralty Res. Lab.(Gt. Brit.) Rep.* No. ARL/RI/84.46/0, 1943.

Elliott, R. Physiological activity and performance: A comparison of kindergarten children with young adults. *Psychol. Monogr.*, 1965, **79**, No. 587.

Frankmann, J. P., & Adams, J. A. Theories of vigilance. *Psychol. Bull.*, 1962, **59**, 257–272.

Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1954.

Haider, M., Spong, P., & Lindsley, D. B. Attention, vigilance, and cortical evoked-potentials in humans. *Science*, 1964, **145**, 180–182.

Hawkes, G. R., Meighan, T. W., & Alluisi, E. A. Vigilance in complex task situations. *J. Psychol.*, 1964, **58**, 223–236.

Hebb, D. O. Drives and the C.N.S. (conceptual nervous system). *Psychol. Rev.*, 1955, **62**, 243–254.

Holland, J. G. Human vigilance. *Science*, 1958, **128**, 61–63.

Jenkins, H. M. Performance on a visual monitoring task as a function of the rate at which signals occur. *Lincoln Lab tech. Rep.* No. 47. Lexington, Mass.: M.I.T. 1953.

Jenkins, H. M. The effect of signal-rate on performance in visual monitoring. *Amer. J. Psychol.*, 1958, **71**, 647–661.

Lindsley, D. B., (Ed.) Radar operator fatigue: the effects of length and repetition of operating periods on efficiency of performance. *Off. Sci. Res. Develpm. Rep.*, No. OSRD-3354. Washington, D.C., 1944.

Luce, R. D. On the possible psychophysical laws. *Psychol. Rev.*, 1959, **66**, 81–95.

McGrath, J. J. Irrelevant stimulation and vigilance performance. In D. W. Buckner & J. J. McGrath (Eds.), *Vigilance: a symposium.* New York: McGraw-Hill, 1963. Pp. 3–19.

Mackworth, N. H. The breakdown of vigilance during prolonged visual search. *Quart. J. exp. Psychol.*, 1948, **1**, 6–21.

Mackworth, N. H. *Researches on the measurement of human performance.* Med. Res. Council (Gt. Brit.), Spec. Rep. Ser. No. 268. London: H.M. Stationary Office, 1950.

Malmo, R. B. Measurement of drive: an unsolved problem in psychology. In M. R. Jones (Ed.), *Nebraska symposium on motivation.* Lincoln Nebr.: Univer. of Nebraska Press, 1958. Pp. 229–265.

Malmo, R. B. Activation: a neuropsychological dimension. *Psychol. Rev.*, 1959, **66**, 367–386.

Malmo, R. B. Activation. In A. J. Bachrach (Ed.), *Experimental foundations of clinical psychology*. New York: Basic Books, 1962. Pp. 386–422.

Malmo, R. B. On central and autonomic nervous system mechanisms in conditioning, learning, and performance. *Canad. J. Psychol.*, 1963, **17**, 1–36.

Nicely, P. E., & Miller, G. A. Some effects of unequal spatial distribution on the detectability of radar targets. *J. exp. Psychol.*, 1957, **53**, 195–198.

Scott, T. H. Literature review of the intellectual effects of perceptual isolation. *Canad. Def. Res. Bd. Rep.*, No. HR66. Ottawa, Ont., 1957.

Smith, R. P. The effects of signal density and signal variability on the efficiency of human vigilance. Unpublished doctoral dissertation, Emory Univer., Atlanta, Ga., 1961.

Stevens, S. S. On the psychophysical law. *Psychol. Rev.*, 1957, **64**, 153–181.

Woodworth, R. S., & Schlosberg, H. *Experimental psychology* (rev. ed.). New York: Holt, 1954.

# Chapter 5

# Facilitation and Interference[1]

WILLIAM F. BATTIG
*University of Maryland*

Throughout the long and somewhat checkered history of psychological research on human learning, a central role has been accorded to transfer effects from previous learning, and to the variety of experimental procedures and phenomena related thereto. However, as shown in Fig. 1, the number of research reports listed by *Psychological Abstracts* that were centrally concerned with transfer has remained relatively stable over the past decade, and has actually declined somewhat over the last as compared with the first half of this period. This picture contrasts sharply with the recent positively accelerated increase in human-learning research in general, and particularly in those paired-associate verbal-learning studies which have traditionally represented the major domain of transfer research (Battig, 1965). As for motor-skills research, which the bibliographies compiled in *Perceptual and Motor Skills* indicate to have remained remarkably consistent in its over-all frequency over the past decade, an even more substantial decline is shown in Fig. 1 for the proportion of the titles in these bibliographies indicative of a direct concern with the study of transfer.

This recent descent of transfer has even reached the point where the very existence of the phenomenon appears to have been ignored by some Skinnerian operant conditioners. Their emphasis upon "reversibility" (Sidman, 1960) and the assumptions underlying this concept represent the direct antithesis of transfer in the sense that subjects purportedly can be returned to "baseline" after learning a particular task, which implies that any transfer

---

[1] Preparation of this chapter was supported in part by Public Health Research Grant HD-01062-01 from the National Institute of Child Health and Human Development. The author is also indebted to several of his present and former students (John K. Berry, Sam C. Brown, Russell B. Johnson, and Philip M. Merikle) for their research and bibliographic contributions to this chapter.

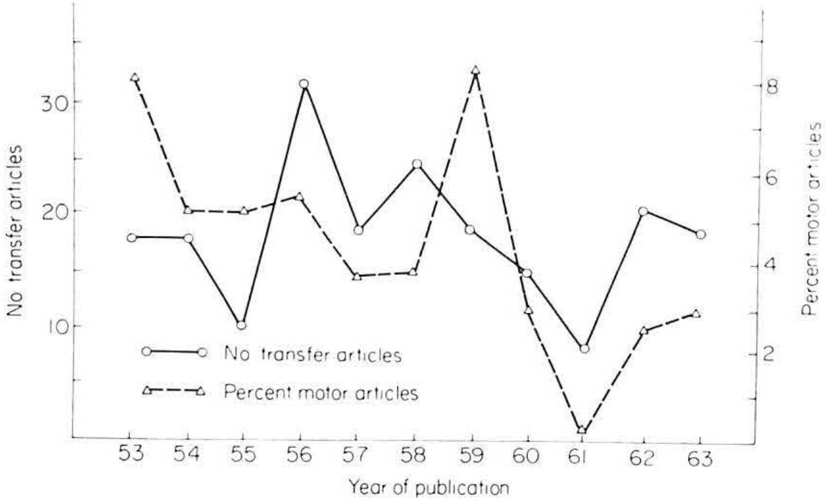


FIG. 1. Number of articles on transfer (titles indicating primary concern with transfer effects on learning listed by *Psychological Abstracts*) and proportion of motor-skills articles directly concerned with transfer (listed in Motor Skills Bibliographies in *Perceptual and Motor Skills*) from 1953 to 1963.

effects from this previous learning are sufficiently unimportant that they can be completely obliterated prior to the subsequent introduction of a new learning task. This "reversibility" assumption, although not put to empirical test by most of those who rely upon it, also represents an essential prerequisite to the closely correlated research strategy of successively introducing a long series of different experimental conditions within a single subject, rather than using different subjects or groups for each experimental task or condition. To the extent that the effects of previous learning are irreversible, i.e., that transfer effects still exist, successively administered variable conditions within a single subject must necessarily become totally confounded therewith.

In point of empirical fact, the status and generality of transfer as a behavioral phenomenon have never before rested on so strong and general an empirical foundation, and even those researchers who choose to ignore it would not quarrel with the statement that previous learning history represents one of the major determinants of performance upon a new learning task. As evidenced by the increasingly prevalent usage of concepts like "learning to learn" or "learning sets," the existence of transfer effects is generally recognized even with respect to learning tasks which are highly dissimilar to one another.

Then why the apparent decline in psychological research explicitly concerned with transfer? Is this a valid indicant of a loss of interest in the study of transfer phenomena? Quite to the contrary, the present discussion will attempt to demonstrate that the experimental methodology formerly associated exclusively with transfer research has now become so widely used throughout other areas of human-learning research that it no longer appears to be a distinct and separate area. An examination of recent issues of the *Journal of Experimental Psychology* and *Journal of Verbal Learning and Verbal Behavior* shows that the majority (61 percent) of those verbal-learning and retention studies *not* containing the term *transfer* or its synonyms in the title had actually included and been directly concerned with relationships between performance on at least two different learning tasks. The currently ongoing increase in this trend is indicated by its rise from 44 percent during the first half of 1963 to 74 percent during the corresponding period of 1964, and reflects primarily a growing concern with the more microscopic analysis of those individual component sources within a single task which produce positive and/or negative effects on performance on a subsequent task, rather than the gross analysis of over-all intertask positive and negative transfer effects. As a consequence, the terms positive and negative transfer have been replaced in the present title by "facilitation" and "interference," respectively, following previous usage of the latter terms to refer to individual component sources of transfer effects as distinguished from over-all transfer (Gagné, Baker, & Foster, 1950).

A somewhat more subtle change in recent experiments employing a transfer methodology has been the use of the second (transfer) task primarily as an indirect means of illuminating the learning processes taking place within the first task, as distinguished from the previously predominant focal interest in second-task performance as it is affected by previous learning of the first task. In addition, the extension of transfer concepts and methods to the investigation of intratask transfer between components of a single multiple-response task has proved to be another potentially significant means whereby transfer and previously separate research areas have been brought closer together. While little of the latter work has as yet extended beyond our own laboratory, it has already yielded some interesting and quite unexpected relationships between facilitation and interference as operative within and between learning tasks. These various ways in which transfer is being assimilated into formerly distinct research areas will be elaborated in the sections to follow.

## Communalities of Transfer with Other Research Areas

The many-faceted use of transfer procedures and concepts in human-learning research explicitly concerned with other types of problems is neither surprising nor new. In particular, studies concerned with the retention phenomena of proactive and retroactive inhibition have necessitated that at least one additional learning task be introduced respectively before and after the task to be tested for retention. As a consequence, it is only recently that a distinction has been maintained with any consistency between these concepts as effects on retention of a previously-learned task, as distinguished from transfer effects on the original acquisition or learning of a subsequent task. Some secondary sources (e.g., Deese, 1958; Kimble & Garmezy, 1963; Morgan, 1961) still fail to make an explicit distinction between so-called proactive effects on retention and transfer effects on learning, a consequence of the necessary inclusion of the experimental paradigm for the latter within that for the former. In the typical proactive-inhibition experiments, however, little attention is directed toward the description and analysis of second-task learning. Moreover, such experimental comparisons as have been made clearly demonstrate that previous learning may affect learning of a second task quite differently from second-task retention (e.g., Greenberg & Underwood, 1950). Consequently, the maintenance of a sharp conceptual and experimental distinction between them is rendered imperative.

But although the recent resurgence in proactive- and retroactive-inhibition research occasioned by recently revised formulations of interference theories of forgetting (Postman, 1961; Underwood & Postman, 1960; Underwood, 1957) has been a major contributor toward the assimilation of transfer with other types of learning research, it has been by no means the only or even the primary source of this integrative development. Experimenters have increasingly come to recognize the economy, as well as the analytic usefulness, of requiring the subject to perform, immediately after the conclusion of the basic acquisition portion of the single-task learning experiment, on one or more related tasks designed to elucidate some specific aspect of the previous learning task or processes involved. Demonstrated to be particularly valuable for such purposes have been "transfer" tasks involving free stimulus or response recall (e.g., Underwood, Runquist, & Schulz, 1959), backward (R-S) learning (e.g., Feldman & Underwood, 1957), or grouping of the items constituting the task (e.g., Fallon & Battig, 1964).

An even more central role has been accorded to intertask facilitation and interference in many of the currently most active verbal-learning research areas, in which the ingenious adaptation of transfer-type designs has constituted a major reason for the current high level of interest. As a prime example, the heated controversy concerning the mechanisms of serial learning owes its very existence to Young's (1959, 1961) original demonstrations of a surprising lack of facilitation from serial to subsequent paired-associate learning of pairs of adjacent items from the serial list, and its maintenance to the several subsequent investigations employing further transfer designs involving related or partially-overlapping serial tasks (e.g., Battig, Brown, & Schild, 1964; Ebenholtz, 1963b; Keppel & Saufley, 1964; Young, 1962), serial lists with variable starting points (Ebenholtz, 1963b; Winnick & Dornbush, 1963), or spatial-position tasks (Ebenholtz, 1963a).

Recent research directed toward the isolation and experimental analysis of component processes involved in paired-associate learning has similarly relied upon transfer designs involving such tasks as free learning of individual items (Underwood *et al.*, 1959), various familiarization procedures (Noble, 1963; Underwood & Schulz, 1960), or verbal-discrimination learning (Battig, Williams, & Williams, 1962). The evaluation of transfer plays an equally significant role in other currently "hot" research areas, such as backward (R-S) associations (Asch & Ebenholtz, 1962), mediation (Horton & Kjeldergaard, 1961; Jenkins, 1963), or verbal satiation (Lambert & Jakobovits, 1960; Yelen & Schulz, 1963). Even those areas not directly involving intertask comparisons, such as short-term memory (Peterson, 1963), clustering or coding processes in free-recall learning (Bousfield, 1953; Cohen, 1963), meaningfulness (Underwood & Schulz, 1960), and one-trial learning (Postman, 1963), are far from devoid of the involvement of transfer effects from other learning. Both the clustering and meaningfulness studies are predicated upon the existence of extensive transfer effects from previously well-learned verbal materials, and several have involved direct efforts to assess these effects experimentally. Likewise, increasing attention is being directed toward the role of type of both prior and interpolated activity in short-term memory studies (e.g., Keppel, 1965), and many of the one-trial-learning studies have involved systematic changes in the materials and conditions of learning which render them virtually indistinguishable from certain types of transfer experiments.

Unfortunately, the use of transfer techniques for the *post hoc* and indirect assessment of component learning processes within a particular task is not without dangers and pitfalls of its own, and current enthusiasm for this undeniably useful type of analytic tool has doubtless served to obscure some of these shortcomings. Particularly questionable is the assumption that the evaluation of performance after a task has been learned provides a valid indicant of the processes involved during the course of its actual learning. It is entirely possible, for example, for performance on a subsequent transfer task to be mediated indirectly through previous learning of a quite distinct variety, as illustrated by a recent experiment in which paired-associate learning of double-function lists was followed by a test of written R-S recall, within which all of the stimulus terms to be recalled necessarily were listed also as response terms on the recall sheet (Battig & Koppenaal, 1965). In this instance, apparent R-S associations could readily be produced by regarding each presented "response" instead as a "stimulus," by recalling its correct S-R associate, and by writing the first item thereof in the space to the left of its recalled response. Equally or perhaps more likely is the possibility that negative effects, resulting from necessary but irrelevant procedural changes from one learning task to another, may be of sufficient magnitude to obscure any facilitation produced by the first-task completion of component learning processes common to both tasks. The puzzling absence of any facilitation whatever from previous familiarization or verbal-discrimination learning of items or pairs subsequently included in a paired-associate task (e.g., Battig *et al.*, 1962; Goss, Nodine, Gregory, Taub, & Kennedy, 1962) may well be due to the latter source of contamination inherent in the use of transfer designs for purposes of intratask component analysis.

## Analysis of Component Processes in Intertask Transfer

In part as a direct consequence of the widespread use of transfer techniques to explicate component processes involved in various types of verbal-learning tasks, there are several recent and highly encouraging signs that the results of such component analyses may be applied with equal profit to the resolution of fundamental problems and contradictions within the area of transfer itself. This significant current trend is most evident with respect to the effects on transfer produced by intertask variations in stimulus and/or response similarity, a problem which probably has received more research and theoretical attention than any other in the area of transfer. What is more, the long-standing resistance of this problem to satisfactory empirical resolution has constituted one of the major deterrents to progress in transfer research. Consequently, the recent history of research on the role of intertask similarity offers a convincing illustration of the gains to be achieved in the understanding of transfer phenomena through the application thereto of the type of analysis of component processes to which transfer methodology has made such significant contributions with respect to single-task acquisition.

Early research on transfer (and retroactive effects on retention) as a function of intertask similarity had led to a number of suggested theoretical and empirical relationships, of which the best known was the Skaggs–Robinson hypothesis (Robinson, 1927) indicating the relationship to be nonmonotonic, such that first-task retention (and presumably also second-task transfer performance) was maximal with high intertask similarity, decreasing to a minimum at moderate similarity levels, and increasing again as complete intertask dissimilarity was approached. Subsequent research accomplished little toward the clarification of this relationship, particularly with respect to resolving (a) the obvious inconsistency in simultaneously attributing both maximal facilitation and interference effects to conditions of high intertask similarity; and (b) the difficulties associated with the specification of either dimensions or amounts of intertask similarity.

Osgood (1949) was responsible for the first significant application of component analysis to what he appropriately termed the "similarity paradox" in transfer and retroaction, through his attempt to describe the separate but interacting effects produced by intertask variations with respect to stimulus and response components. His proposed resolution took the form of a three-dimensional "transfer surface," which indicated a reversal from over-all positive to negative transfer at some unspecified intermediate point of response similarity, and an increased magnitude of transfer in either direction with increased amounts of stimulus similarity.

Inasmuch as Osgood's transfer surface was admittedly based upon considerably less than adequate empirical data, several subsequent investigators attempted to test more precisely its adequacy and to fill in the gaps with respect to similarity conditions where empirical data were sparse (e.g., Bugelski & Cadwallader, 1956; Dallett, 1962; Long & Donahoe, 1962; Wimer, 1964; Young & Underwood, 1954). Instead of clarifying the issue, these studies have shown little consistency either with Osgood's analysis or with one another, the results ranging from essential agreement with Osgood's transfer surface (Dallett, 1962) to marked discrepancies from it (Bugelski & Cadwallader, 1956). Unfortunately, little has been accomplished toward the identification of the major factors responsible for these inconsistencies, although the multidimensionality of similarity and lack of satisfactory scaling metrics thereof loom as the most probable sources of difficulty. Such problems are especially pronounced with respect to the "dissimilar" half of Osgood's response-similarity continuum proceeding from "neutral" to "opposed" and "antagonistic." Since there is considerable question as to whether "opposition" is less similar psychologically than "neutrality," and Osgood employed the neutral-opposed portion of the continuum only for response and not for stimulus similarity, much of the confusion disappears if neutrality rather than opposition is considered as the point of minimal similarity. Suggested as additional sources of those empirical discrepancies that remain have been such factors as (a) whether similarity is defined in terms of semantic or meaningful relationships, or number of overlapping elements (Underwood, 1961); (b) failure to control for differences in list difficulty (Underwood, 1961; Dallett, 1962); and (c) possible differences between intertask similarity effects on second-task learning and first-task retention (Dallett, 1962).

Clearly the resolution of this "similarity paradox" awaits the development of several necessary methodological prerequisites, and the concentration on these problems of scaling and variable specification appears more profitable at present than continued efforts directed toward precise quantitative specification of intertask similarity effects on transfer without adequate tools for the job. Little beyond added confusion is contributed to the issue by such as Houston's (1964) reincarnation of the Osgood (1949) transfer surface to encompass conditions where first-task stimuli become second-task responses and vice versa. Not only is Houston's surface based on even less empirical evidence than was available to Osgood, but it is contradicted by such evidence as is available (e.g., Murdock, 1958).

For these and other reasons, the majority of recent experiments on intertask similarity effects in transfer and retention have been limited to the various combinations of identity and neutrality with respect to constituent stimulus and/or response terms across first and second lists. If the first list is designated as consisting of A-B pairs, the second lists for the resulting transfer paradigms are commonly labeled as follows: (1) *A-C paradigm*, first-list stimuli paired with new responses; (2) *C-B paradigm*, first-list responses paired with new stimuli; (3) *A-Br paradigm*, first-list stimuli and responses re-paired differently from the first list; and (4) *C-D paradigm*, new stimuli paired with new responses. The latter C-D paradigm, which represents conditions of minimal similarity of both stimulus and response terms, thus provides the appropriate control baseline for the evaluation of specific transfer effects attributable to the specific similarity relationships within the other paradigms, free of contamination by nonspecific sources of transfer (e.g., learning how to learn).

Principally on the basis of Bruce's (1933) classic study, it had long been widely accepted that over-all negative transfer results from old stimuli paired with new responses (A-C paradigm), whereas new stimuli paired with old responses (C-B paradigm) produce positive transfer. The explicit postulation of maximum negative transfer under the A-Br paradigm did not come until the Gagné *et al.* (1950) analysis uniquely concerned with similarity effects on transfer in discriminative motor tasks. However, corroboration of the finding of greater A-Br than A-C interference in those verbal paired-associate tasks typifying the bulk of research in this area was not long in coming (Porter & Duncan, 1953) or in being further substantiated with more adequate controls (e.g., Besch & Reynolds, 1958; Twedt & Underwood, 1959).

Sharply contrasting with the results summarized above have been Mandler's findings of A-Br facilitation of a magnitude even greater than that found for the C-B paradigm with a verbal task (Mandler & Heinemann, 1956), and of maximal (although not entirely consistent) interference effects under the A-C paradigm with a verbal as well as with a motor task (Mandler, 1954). Like the earlier analyses, Mandler has consistently reported significant C-B facilitation, but a number of recent investigators have demonstrated instead significant interference under this paradigm (e.g., Harcum, 1953; Jung, 1962; Kausler & Kanoti, 1963; Postman, 1962; Twedt & Underwood, 1959). Several hypothesized sources of these discrepancies have subsequently been investigated, including mixed vs. unmixed lists (Twedt & Underwood, 1959), and degree of first-list learning (Jung, 1962; Mandler & Heinemann, 1956; Postman, 1962), but only meaningfulness of stimulus and/or response terms has been shown consistently to interact significantly with relative performance under the various transfer paradigms (Jung, 1963; Merikle & Battig, 1963; Merikle, 1964).

It has been largely through these manipulations of item meaningfulness that the need for further analysis of the differential transfer effects produced by various component processes within the paired-associate task has come to the fore as a means of resolving the discrepancies arising from previous research. Although there is as yet little agreement as to either the number or nature of these component paired-associate processes relevant to comparisons between transfer paradigms, research attention to date has focused primarily on a distinction between response-learning and associative stages (Underwood & Schulz, 1960), with the latter being further subdivided into forward, backward, and contextual associations of an unspecified nature (Keppel & Underwood, 1962; McGovern, 1964).

Because this approach to the analysis of transfer phenomena is relatively new, and much additional work needs to be done before the numerous and complex issues raised thereby can be satisfactorily resolved, we shall attempt here merely to illustrate the potential effectiveness of the approach, and to suggest some general features of its probable theoretical resolution. Merikle's (1964) study has been selected for purposes of the present illustration, because it alone has systematically varied both stimulus and response meaningfulness under all four of the basic paradigms (A-C, C-B, A-Br, and C-D), and thus goes somewhat beyond other similar empirical (Jung, 1963) and theoretical (Martin, 1963) analyses in offering evidence relevant to most if not all of the component processes presumed to enter into the paired-associate learning and transfer effects to be accounted for.

In Merikle's experiment, 16 independent groups of eight subjects each represented all of the four basic transfer paradigms within each of four meaningfulness (M) conditions consisting of all possible combinations of stimulus and response terms at high (H) and low (L) meaningfulness levels, *cvc* trigrams of 85–94 percent and 10–20 percent (Archer, 1960) association value, respectively. In an attempt to reduce interpair differences in first-list learning, Merikle employed an adjusted-learning recall procedure whereby each pair was removed from the list as soon as it had elicited two consecutive correct recalls, with a final trial on the entire list after all pairs had reached this criterion. Although this procedure of dropping out individual pairs upon attainment of a specified performance criterion had characterized much of the earlier research (e.g., Osgood, 1946; Bugelski & Cadwallader, 1956), it has subsequently fallen into disuse if not disrepute in experiments concerned with transfer under the various paradigms, as well as in other areas of transfer and retention research. However, a subsequent direct comparison between adjusted-learning and standard paired-associate recall procedures across the four transfer paradigms has demonstrated the transfer effects to be closely comparable under the two procedures (Merikle, Manatis, & Battig, 1965).[2]

As shown in Table I, both direction and magnitude of transfer under the various paradigms differed markedly and differentially under Merikle's various M conditions, relative to the appropriate C-D (control) paradigm. Increasing response M served to eliminate over-all facilitation under the A-Br and C-B paradigms, and produced significant interference under the A-Br and A-C paradigms. Increased stimulus M, on the other hand, led to a greater magnitude of both facilitation and interference, depending on whether response M was L or H. Not only are differential types of transfer effects produced by variations in stimulus and response M, but these appear to interact substantially with one another in determining over-all direction and amount of transfer.

---

[2] It is most unfortunate, in the author's opinion, that the standard paired-associate anticipation procedure, wherein the entire list of pairs is presented on every trial, has so completely dominated investigations of paired-associate learning and transfer in most laboratories. This has been true even of investigations directly concerned with the manipulation of degree of original learning, despite the fact that such complete-list or noncorrection procedures inevitably lead to substantial differences between individual pairs in degree of original learning, besides preventing either effective evaluation or control of this variable. Since Merikle's (1964) results still indicated better learning of those pairs which were first learned and removed from the list, we are currently attempting to specify·and equate degree of learning of individual items through a "correction" procedure, whereby the number of presentations and correct responses required for each item are both increased as a function of number of errors made upon that item. Such procedures, of course, must otherwise follow the basic recall procedure (Battig & Brackett, 1961), since the necessary presentation under the anticipation procedure of each S-R pair immediately following its attempted anticipation offers the opportunity for additional learning which cannot be reflected in the subject's performance until the next trial, and thus renders impossible the direct measurement and control of level of learning even for the entire list, much less for the individual pair.

TABLE I

MEAN SECOND-LIST TOTAL ERRORS AND PERCENT TRANSFER FOR EACH GROUP[a]

| *Transfer paradigm* | | *H-H* | *L-H* | *H-L* | *L-L* | *Total* |
|---|---|---|---|---|---|---|
| A-Br | Errors | 11.00 | 20.75 | 23.88 | 31.50 | 21.78 |
| | % T | −120.0 | −15.3 | 34.1 | 33.2 | |
| A-C | Errors | 13.13 | 20.25 | 32.88 | 46.50 | 28.19 |
| | % T | −162.6 | −12.5 | 9.3 | 1.4 | |
| C-B | Errors | 5.00 | 16.63 | 14.00 | 34.13 | 17.44 |
| | % T | 0.0 | 7.6 | 61.4 | 27.6 | |
| C-D | Errors | 5.00 | 18.00 | 36.25 | 47.13 | 26.60 |

[a] Merikle (1964).

Also illustrated by Merikle's experiment is the potential value of giving a series of recall and associative-matching tests at the conclusion of second-list learning, which in this instance substantially clarified the nature of the interactive effects of stimulus and response M on transfer under the various paradigms. When the subject was presented with all stimulus terms and asked to give both first- and second-list responses to these, first-list response recall was markedly reduced with L stimuli, relative both to second-list responses with L stimuli and to recall of both first- and second-list responses with H stimuli. This effect was particularly pronounced under the A-Br paradigm, but existed to a lesser extent also for A-C and C-D paradigms, and thus suggests "unlearning" of first-list associations to be maximal with L stimuli. By contrast, first-list associations with H stimuli instead tend to be maintained and possibly even used as mediators during second-list learning (Postman, 1961), thus accounting for the consistently greater magnitude of over-all transfer with H than with L stimuli. Since L-stimulus recall was also considerably below that for H stimuli, there appears to be a greater tendency toward stimulus selection of a part of the nominal stimulus to serve as the functional stimulus with L than H stimuli (Underwood, 1963). The reduced interlist transfer effects with L stimuli may thus reflect the increased tendency for the subject to select a functional stimulus for the second list which is different from that of the first list.

Merikle's over-all differences between paradigms, and the effects of response M upon these differences, can readily be interpreted in terms of the previously established mechanisms of (a) facilitation of second-list response learning under A-Br and C-B paradigms, especially with L responses; (b) interference due to changed forward associations under A-Br and A-C paradigms; and (c) a lesser interference due to changed backward associations under A-Br and C-B paradigms. However, the elimination of over-all interference with L responses under the A-C paradigm is indicative also of a stimulus-learning process, which is enhanced by the increasing number of first-list presentations required to reach criterion with L as compared with H responses. And since consistently more negative (or less positive) effects were produced when stimuli and responses were alike in M (L-L and H-H) than when they differed (L-H and H-L), a lack of differentiation between stimulus and response terms is also implicated as a possible interference-producing factor in this experiment.

Summing up, a complete interpretation of Merikle's results indicates the operation of learning processes involving (1) forward (S-R) associations; (2) backward (R-S) associations; (3) response learning; (4) stimulus learning; (5) stimulus selection; (6) differentiation of stimuli from responses; and (7) interlist mediated associations. But what is more impressive than the sheer number of these processes is their consistency with processes indicated to be involved in paired-associate learning by previous research, much of which has made use of transfer-type designs. In view of the extensive contributions made by transfer procedures toward the understanding of intratask learning, it seems entirely appropriate that the latter should now be applied to the analysis of intertask transfer, even though the present example serves merely to illustrate an extremely promising experimental approach to the many unresolved questions concerning the role of these various learning processes in transfer situations.

## Intratask Facilitation and Interference

There is yet another important way in which single-task learning research has been influenced by transfer, through the use of transfer-type concepts to account for intratask similarity effects in multiple-response learning tasks. This type of approach dates back at least to Gibson's (1940) generalization-differentiation theory, recently analyzed and criticized by Underwood (1961), which has been extensively applied to learning tasks involving high levels of similarity between intratask stimuli, responses, or both. It has been widely assumed that any intratask transfer effects can be eliminated if the various intratask components are made maximally dissimilar to one another, as has been typical of intertask transfer research.

Recent research in our laboratory, however, has demonstrated the operation of such intratask transfer effects in paired-associate lists even at very low similarity levels, and has indicated a major portion of paired-associate learning to consist of the overcoming of interpair interference produced by the context of other pairs within the list. Depending particularly upon the state or degree of learning of other pairs within the list, these interpair effects may either facilitate or interfere with the learning of any individual pair. Significant interpair interference is likely to result from a context of new or previously incorrect pairs, reaching a maximum for unlearned pairs with a low to moderate degree of prior practice (Brown, 1964). However, learning of a list of new pairs will be significantly facilitated by the inclusion of a few well-learned pairs from a previous list, even if these are added to rather than substituted for other pairs in the list (Brown & Battig, 1966). That these effects have been demonstrated, even when the pairs are minimally similar to one another, offers convincing evidence for the generality of such interpair relationships in paired-associate tasks, as well as indicating the potential usefulness for the investigation of these intratask transfer effects of concepts and experimental procedures developed within intertask transfer research. Likewise, as was demonstrated in the preceding section, such intratask analyses can be expected to contribute also toward the increased understanding of intertask transfer effects.

Despite the paucity of such intratask transfer research, our limited use of this approach has already suggested a potentially very significant if somewhat paradoxical principle relating intratask to intertask transfer, that *intertask facilitation is produced by intratask interference*. That is to say, if learning of a first task is carried out under conditions of high intratask interference, this is likely to result in maximal facilitation of the subsequent learning of a partially similar or related second task. Such a principle was originally formulated and developed by Russell Johnson in an unpublished Ph.D. dissertation, which still constitutes its most impressive single documentation (Johnson, 1964). Its relevance to the present discussion is further enhanced by its direct concern with a problem which once had been one of the most important in the area of intertask transfer. This area of *stimulus predifferentiation* concerns the establishment of the existence of and conditions determining facilitation of second-task learning under a special variety of A-C transfer paradigm, wherein first-task responses are maximally dissimilar both to one another and to second-task responses. Stimulus predifferentiation research, however, has been plagued by a plethora of small and inconsistent effects (Arnoult, 1957), so that only a few current investigators have maintained an interest therein (e.g., Ellis & Muller, 1964; Vanderplas, Sanderson, & Vanderplas, 1964).

In fact, Johnson's (1964) experiment was initially dictated by his own previous (1962) contribution to the inconsistencies in stimulus-predifferentiation research, wherein he had attempted to include in a single experimental design not only typical stimulus-predifferentiation conditions, but also the opposite *stimulus-equivalence* paradigm (Miller & Dollard, 1941), under which the pairing of identical or similar responses with initially distinct stimuli is alleged to decrease the distinctiveness or differentiation between them. Instead of finding greater facilitation on a subsequent multiple choice recognition task following the pairing of similar stimuli (nonsense shapes) with distinct unrelated verbal labels, however, Johnson's (1962) results had shown recognition accuracy to be significantly better for dissimilar shapes previously paired with similar categorically related response labels, even though the recognition task was identical under the two conditions. That is to say, so-called stimulus predifferentiation training had actually produced less predifferentiation than had stimulus equivalence pretraining.

Johnson's subsequent (1964) study thus derived from a vaguely defined notion that these previous results might somehow have been attributable to the maximal congruence or "compatibility" of stimulus, response, and frequency components of the pretraining task, since all shapes constituting each of the three "classes" of similar shapes had always been presented the same number of times during the pretraining task, as also had all of the response labels within each categorically related class. Consequently, as illus-

TABLE II
PRETRAINING CONDITIONS USED BY JOHNSON (1964)[a]

| *Label* | *C-SLF* | *C-SL* | *C-SF* | *C-LF* | *C-I* | *D-LF* | *D-I* |
|---|---|---|---|---|---|---|---|
| *Marine* | $A_1 - 4$ | $A_1 - 4$ | $A_1 - 4$ | $A_1 - 4$ | $A_1 - 4$ | $A_1 - 4$ | $A_1 - 4$ |
| Army | $A_2 - 4$ | $A_2 - 8$ | $B_1 - 8$ | $B_1 - 4$ | $B_1 - 8$ | $B_1 - 4$ | $B_1 - 8$ |
| Navy | $A_3 - 4$ | $A_3 - 16$ | $C_1 - 16$ | $C_1 - 4$ | $C_1 - 16$ | $C_1 - 4$ | $C_1 - 16$ |
| *Feminine* | $B_1 - 8$ | $B_1 - 4$ | $A_2 - 4$ | $A_2 - 8$ | $A_2 - 8$ | $D_1 - 8$ | $D_1 - 4$ |
| Masculine | $B_2 - 8$ | $B_2 - 8$ | $B_2 - 8$ | $B_2 - 8$ | $B_2 - 16$ | $E_1 - 8$ | $E_1 - 8$ |
| Neuter | $B_3 - 8$ | $B_3 - 16$ | $C_2 - 16$ | $C_2 - 8$ | $C_2 - 4$ | $F_1 - 8$ | $F_1 - 16$ |
| *Ampere* | $C_1 - 16$ | $C_1 - 4$ | $A_3 - 4$ | $A_3 - 16$ | $A_3 - 16$ | $G_1 - 16$ | $G_1 - 4$ |
| Volt | $C_2 - 16$ | $C_2 - 8$ | $B_3 - 8$ | $B_3 - 16$ | $B_3 - 4$ | $H_1 - 16$ | $H_1 - 8$ |
| Ohm | $C_3 - 16$ | $C_3 - 16$ | $C_3 - 16$ | $C_3 - 16$ | $C_3 - 8$ | $I_1 - 16$ | $I_1 - 16$ |

[a] Similar (class) shapes are designated by the same letter with different subscripts, whereas different letters represent dissimilar shapes. The numbers (4, 8, 16) designate frequency of pretraining presentation for each shape-label pair.

trated in Table II, all possible combinations of congruence-incongruence of the shape, label, and frequency components of the pretraining task were included, using nine dissimilar (D) shapes as well as three classes each of three highly similar (C) shapes. Under each of the resulting seven congruence conditions, the nine stimulus shapes were paired for different subgroups of nine subjects each with either (a) three distinct response labels (the ones italicized in Table II), each being paired with three of the shapes; or (b) all nine of the labels given in Table II, which represented three of Cohen's (1963) maximally interrelated three-word exhaustive categories.

For present purposes, it will greatly simplify the exposition of the essential features of Johnson's enormously complicated results if we ignore the frequency component of the pretraining task, since the effects produced thereby are largely irrelevant to the present discussion (although the opportunity provided to "group" together subsets of pretraining pairs on the basis of differential frequency produced extremely interesting and potentially important effects both on pretraining and recognition performance). Furthermore, the remaining pretraining conditions can most meaningfully be considered as reflecting systematic variations in amount of intratask interference, as indicated by the differential levels of performance attained on the pretraining task. More specifically, these results clearly demonstrated fewer correct pretraining responses and greater intratask interference to be produced with (a) similar class (C) than dissimilar nonclass (D) shapes; (b) class-label incongruence (C-SF, C-LF, and C-I conditions) than congruence (C-SLF and C-SL conditions); and (c) nine than three different response labels. These pretraining performance differences, incidentally, are in close agreement with previous findings based on similar types of intratask variations in classical types of verbal-learning tasks (e.g., Richardson, 1958; Underwood & Schulz, 1961, Experiment III).

TABLE III

MEAN PAIRED-ASSOCIATE PRETRAINING AND RECOGNITION PERFORMANCE[a]

| *Congruence condition* | *No. of labels* | *Paired-associate (mean % correct)* | *Recognition accuracy (%)* |
|---|---|---|---|
| C-SLF and C-SL | 3 | 92.9 | 31.9 |
| D-LF and D-I | 3 and 9 | 76.8 | 34.1 |
| C-SLF and C-SL | 9 | 67.1 | 36.3 |
| C-LF, C-SF, and C-I | 3 | 59.9 | 36.9 |
| C-LF, C-SF, and C-I | 9 | 54.4 | 42.2 |

[a] Johnson (1964).

Consequently, Johnson's pretraining conditions can be roughly ordered as in Table III from lesser to greater amounts of intratask interference, as defined either by the actual pretraining performance measures presented in the table, or in terms of the differential contributions of the stimulus, congruence, and response components as outlined above. The important finding for present purposes, however, is the consistent inverse relationship between pretraining performance and accuracy of subsequent recognition of the pretraining stimuli, as given in the last column of Table III. It is clear that the poorer the pretraining performance, the greater the facilitation of recognition accuracy. Or, as stated earlier, intertask facilitation was produced by (or at least closely associated with) increasing intratask interference.

To the extent that recognition performance is a satisfactory measure of amount of stimulus learning, Johnson's results clearly demonstrate this to be systematically increased under conditions of greater intratask interference. However, he was able to account for his results more precisely and elaborately on the basis of the somewhat different types of additional stimulus learning produced by stimulus, response, and associative sources of intratask interference. Both stimulus interference (produced by increasing stimulus similarity) and associative interference (produced by incongruent shape-label pairings such that similar shapes are paired with dissimilar labels and vice versa) can be assumed to produce learning primarily of unique or differentiating cues for each stimulus, that is, cues which characterize that stimulus alone and not any of the other stimuli. While the nature of this emphasis on unique cues is similar for stimulus and associative interference, the present evidence indicates it to be more pronounced for the latter source. As for response interference (resulting from the increase from three to nine response labels), this can be assumed to produce learning of those stimulus cues which are most readily associated with the response label, irrespective of whether these cues are common to two or more stimuli, or unique to an individual stimulus. Evidence supporting the latter contention is provided by Johnson's further finding of significantly superior recognition accuracy for nine- (41.8 percent) over three-label conditions (33.4 percent) only if the subject was informed as to how many of the shapes in each multiple choice recognition set had actually appeared in the pretraining task, whereas virtually identical accuracy was found for nine (32.2 percent) and three labels (32.3 percent) when the subjects were free to select as many or as few shapes as they wished. Thus without added information as to how many shapes are to be chosen, the additional learning of common nondifferential as well as unique cues, as is produced by increasing response interference, appears to obscure any tendency toward increased recognition accuracy, the latter becoming evident only when the influence of these common cues is reduced by telling the subject how many correct shapes are present.

Quite contrary to any of the previous conceptions concerning the kind of pretraining conditions which maximize stimulus predifferentiation, Johnson's results indicate these to involve maximal intratask interference resulting not only from the use of high levels of stimulus similarity, but also associative interference produced by maximal incongruence of pairings of these stimuli with response labels, particularly if the latter are sufficiently similar to produce maximal response interference. Many of the problems and inconsistencies which have plagued previous research concerned with stimulus predifferentiation already appear to be resolvable through this new conceptualization predicated upon differences in intratask interference. Moreover, through its provision of conditions reliably yielding substantial amounts of stimulus predifferentiation, this approach for the first time has made the phenomenon itself amenable to effective experimental analysis.

Johnson's results and interpretations alone, of course, do not permit widespread generalization beyond his particular stimulus-recognition transfer task. However, other results from our laboratory are indicative of a similar facilitative effect from previous intratask interference within a serial-learning type of transfer task, wherein it has been shown that significantly more second-list facilitation is produced by the inclusion therein of items taken from the middle rather than from the beginning of a previously learned serial list (Battig, Brown, & Schild, 1964). More recently, S. C. Brown (unpublished work, 1965) has found a similar increase in facilitation for items from the middle as compared with the ends of the list with respect to performance upon them within a subsequent paired-associate task.

Research which is still in progress further suggests that the widely disparate effects produced by various individual-item pretraining procedures upon subsequent paired-associate performance on tasks containing these pretrained items, which have run the entire gamut from negative to high positive transfer, may also be attributable at least in part to differences between the pretraining tasks with respect to intratask interference. Such an explanation appears particularly applicable to the marked discrepancy between the sizable paired-associate facilitation produced by a multiple choice recognition task, wherein each correct item must be selected from within a set that includes also several highly similar alternatives (Battig, Merikle, & Schild, 1965; Thompson, 1965), as contrasted with findings of either interference or no effect if the items have appeared on a prior two-alternative verbal-discrimination task (Battig & Brackett, 1963; Battig *et al.*, 1962), or under single-item familiarization procedures (e.g., Goss *et al.*, 1962). However, during 1965 data collection was completed in an experiment explicitly designed to assess the role of pretraining intratask interference upon subsequent paired-associative transfer. Preliminary analyses of these data indicate paired-associate facilitation on pretrained relative to new items to be limited to conditions of maximal pretraining intratask interference, in support of the proposed explanation. However, this experiment also reveals no facilitation whatever under the pretraining conditions most closely corresponding to those which had produced it maximally in previous research.[3] Thus while available evidence offers encouraging indications that pretraining intratask interference may be a significant source of subsequent paired-associate facilitation, a satisfactory evaluation and understanding of the effects on paired-associate transfer produced by variations in pretraining intratask interference is still many experiments away.

To the extent that the suggested inverse relation between intra- and intertask transfer effects seems contradictory, it is principally because of the predominant emphasis upon the production of maximal efficiency of original learning, and the widespread assumption, not without empirical support (e.g., Underwood, 1964), that degree of original learning is the primary if not the sole determinant of retention and presumably also of transfer. However, it appears to be not at all inconsistent with the currently prevalent interference explanation for both retention and transfer (e.g., Postman, 1961) that these should be facilitated if original learning has taken place under conditions involving the same kinds of interference that are normally responsible for the subsequent loss of effects of this original learning. From this statement it appears obvious that facilitation of subsequent performance should be increased if original learning has taken place under conditions of high intratask interference, provided that this interference is similar to that which is responsible for forgetting. Maximally efficient learning conditions, whatever facilitation they may produce in original learning, are virtually worthless if they also eventuate in minimal facilitation of subsequent transfer and retention, as the present analysis clearly suggests to be the case. To the extent that facilitated transfer and retention are demonstrated by future research to result from original learning under conditions of extensive intratask interference, an entirely new approach to education and other applied learning situations may be called for. As far as the present discussion is concerned, however, it is sufficient to note the direct derivation of this suggested relationship from what has thus far been at best the exploratory investigation of intratask sources of facilitation and interference as these are related to intertask transfer effects.

---

[3] Subsequent completion of the analysis of this experiment, along with another performed subsequent thereto, produced several further lines of evidence consistent with the proposed principle, to be reported in detail elsewhere. These analyses also point to several additional factors (type and length of item, learning of incorrect and irrelevant correct alternatives during pretraining, and proportion of pretrained pairs included in a mixed paired-associate list) significantly related to the existence and amount of paired-associate facilitation resulting from variations in pretraining intratask interference.

It should be obvious from the preceding discussion that future research on this problem will require extensive efforts directed toward analyzing and distinguishing between various component sources of intratask interference, consisting of more thoroughgoing analyses along the lines suggested by Johnson's (1964) relatively gross distinctions among stimulus, associative, and response sources of intratask interference as described herein. So this section also is appropriately concluded by pointing up the essential role to be accorded to intratask component analysis in future transfer research.

## Implications for Motor-Skills Transfer Research

Since the preceding discussion has been based almost entirely upon verbal types of transfer tasks, any extension of the ideas suggested therein to other types of learning situations must proceed with considerable caution. As far as motor-skills transfer is concerned, the proposal in the preceding section that intratask interference leads to intertask facilitation may appear to be inconsistent with the author's earlier conclusion that facilitation produced by relevant verbal pretraining is progressively eliminated as complexity of the motor transfer task is increased (Battig, 1956). This inconsistency is an illusory one, however, since the present principle concerns the degree of intratask interference present in the first task as it affects subsequent performance on the second, whereas the previous analysis was limited to the effects produced by variations in complexity (and presumably although not necessarily also intratask interference) involving primarily the second task. It is entirely possible, and not at all inconsistent with the research previously described as bearing on the principle, that the degree of intertask facilitation produced by intratask interference within the first task will systematically decrease with increasing intratask interference within the second task. Although the latter interpretation is presently sheer speculation, it clearly merits direct experimental investigation with respect to both verbal and motor types of transfer tasks.

That the present analysis is not devoid of relevance to motor-skills transfer is nicely demonstrated by its ability to account for a puzzling and previously uninterpretable finding in the aforementioned Battig (1956) experiment, which had been accorded the status of an artifact for nearly a decade. This finding was the significant quadratic component in the curve depicted in Fig. 2 displaying amount of transfer as a function of complexity of the motor-transfer task following relevant-S pretraining, such that the general trend toward decreasing positive transfer with increasing complexity was reversed for the motor task of the highest (four-finger) complexity. This relevant-S pretraining task had required the subject to pronounce nonsense words constituted by the sequence of individual letters adjacent to each individual light for several stimulus-light patterns, which were subsequently to be matched by finger movements in the finger-positioning transfer task. Since the number of stimulus lights in the patterns for each subject corresponded to the number of fingers to be used in the finger-positioning task, this necessarily meant that length of the pretraining "word" increased from one to four letters as motor-task complexity increased. Furthermore, the letters were arranged so that only the second letter was a vowel, with the result that the pretraining responses were all readily pronounceable except under the most complex four-finger condition, wherein the required pronunciations included such as QAPB and XENF. Clearly there was a substantial discontinuity between the three- and four-finger relevant-S pretraining conditions with respect to difficulty of pronunciation of the required responses, enhanced further by the consistent consonant-vowel-consonant arrangement for the three-finger group which probably provided for as easy or easier pronunciation than any of the other complexity conditions.

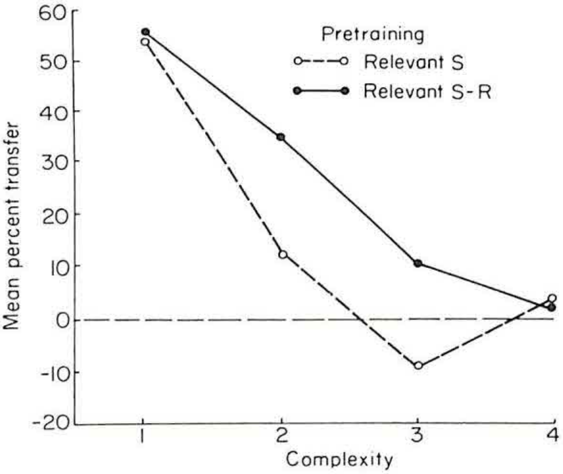


FIG. 2. Percent transfer from relevant-S and relevant-S-R types of verbal pretraining to subsequent motor performance on a finger-positioning task as a function of motor-task complexity defined in terms of number of fingers (1–4) used in the finger-positioning task (Battig, 1956).

Assuming that increased difficulty of pronunciation had represented a source of increasing intratask interference within the four-finger relevant-S pretraining task, the finding that this had produced an increased facilitative effect would follow directly from the present postulation of increased intertask facilitation due to increased intratask interference in the first task. Also consistent with this principle is the absence of a similar reversal in the relevant-S-R curves of Fig. 2, since this pretraining task required only that the subject describe the direction and magnitude of the finger movement or movements necessary to match each stimulus pattern, and thus would entail little or no change in the amount of intratask interference as complexity of the relevant-S-R pretraining task was increased.

Just as in the stimulus predifferentiation studies alluded to in the previous section, several of which have made use of motor transfer tasks, the limited facilitative effects typically resulting from various types of relevant verbal pretraining, even when there are obvious interrelationships between the associations required on the pretraining and motor transfer tasks, may well reflect the minimal pretraining intratask interference typically involved under these conditions (e.g., Battig, 1954; Gagné & Baker, 1950; Goss & Greenfield, 1958; McAllister, 1953).

Comparisons of transfer as affected by variations in first-task difficulty, however, have shown no marked or consistent tendency toward greater second-task facilitation with increases in first-task difficulty (Day, 1956; Holding, 1962), a finding seemingly inconsistent with the postulation of intratask interference as a source of intertask facilitation. Task difficulty, of course, encompasses a diverse assortment of incompletely specified dimensions of variation, and can be increased in many ways without substantially varying the degree to which components of the task interfere with one another. A critical examination of the reported studies reveals that few have involved clearcut variations primarily in intratask interference, and fewer still have included orthogonal manipulations in difficulty of both the first and second tasks, as is necessary in order to avoid confounding of first-task difficulty with intertask similarity. Thus while the present formulation receives little support from studies of motor transfer as a function of first-task difficulty, neither is it explicitly contradicted by them. Moreover, the present attempt to delineate a specific source of intertask transfer effects as a function of one type of variation related to task difficulty offers an example of the level of analysis in terms of individual task components advocated both by Day (1956) and Holding (1962) as the road to resolution of the inconsistencies plaguing research on the difficulty variable in transfer. In any event, the obvious and direct test of the applicability of the present formulation to motor-skills transfer remains to be carried out, and will require independent systematic manipulation of the amount of intratask interference both within the first and second tasks, with minimal contamination by other dimensions of task variation unrelated to intratask interference.

There is, however, another puzzling result emanating from studies of motor-skills transfer and retention that appears entirely consistent with the present postulation of intratask interference as a prime source of intertask facilitation. Bilodeau and Bilodeau (1961) have recently reviewed and commented upon the surprisingly small and temporary interference effects produced even by complete reversals of associative relationships between stimuli and discriminative motor responses throughout the extensive research program by Don Lewis and his associates (e.g., Barch & Lewis, 1954; Lewis, McAllister, & Adams, 1951), concluding that, "it is not easy to produce interference" (Bilodeau & Bilodeau, 1961, p. 262). A common characteristic of the motor tasks predominantly employed in these investigations (e.g., the Complex Coordination Test and Star Discrimeter) has been the simultaneous presence of several different stimulus-response pairings, and of higher than usual levels of similarity particularly between the responses entering into the various pairings. Both of these factors would be expected to produce levels of intratask interference which are not only of considerable magnitude, but also are attributable to very similar sources across the original and transfer learning tasks. According to the present analysis, these are precisely the conditions that would be expected to produce maximal intertask facilitation, which appears to be sufficient to virtually cancel out those specific sources of interference resulting from changed associative relationships between stimuli and responses across the two tasks.

Finally, a few speculations may be in order concerning the possibilities of extending the use of transfer techniques for purposes of intratask component analysis of motor tasks, as has been done so extensively in verbal-learning situations. Recognition of the importance of distinguishing between the component skills involved in complex motor performance dates back at least as far in studies of motor as in verbal learning; it was vociferously advocated by K. U. Smith and his associates for well over a decade (e.g., Davis, Wehrkamp, & Smith, 1951; Rubin & Smith, 1952), and became a central focus of Adams' (1964) recent review of the motor-skills literature. However, Smith's approach to the component analysis of motor tasks, as well as that of subsequent investigators (e.g., Archer & Bourne, 1956; Archer, 1958), has involved attempts to measure simultaneously and independently the separate components of the complex overall motor skill.

As a result, the components of motor-learning tasks have tended to be distinguished primarily on the basis of physical properties and/or convenience of measurement, as illustrated by the major distinction made by

Smith *et al.* between "manipulation" and "travel" components on the basis of whether or not the subject's hand is in contact with one of the sequence of knobs or switches to be operated in performing the task (Davis *et al.*, 1951; Harris & Smith, 1953). The distinctions to be made among task components thus rest upon the application of correlational techniques rather than direct experimental manipulations, and a lack of correlation between measures across trials is adduced as evidence for the separability of the task components measured thereby. The validity of such an approach is thus predicated upon the degree to which these physically distinguishable components represent corresponding distinctions between behavioral processes. As a consequence, serious questions are raised by motion-picture analyses indicating a transference of a substantial portion of the "manipulative" component to the noncontact portion which is being measured as "travel," which may well account for the commonly observed lack of correlation between manipulation and travel time measures across successive practice trials (Annett, Golby, & Kay, 1958).

Inasmuch as the level of research activity currently directed toward the analysis of component processes in complex motor tasks compares quite unfavorably with the current status of verbal-learning research, it is tempting to suggest the failure to adopt intertask transfer techniques for this purpose as the primary cause of the discrepancy. To be sure, some of the stimulus-predifferentiation and verbal-pretraining studies alluded to above can be considered as examples of such a transfer approach, as also can those occasional attempts to investigate transfer effects produced by previous practice on relatively gross sensory and/or motor components of the transfer task (e.g., Ammons, 1951; Battig, 1954). However, these studies have for the most part been secondarily concerned with intratask component analysis of the motor task, and have contributed relatively little to the understanding thereof.

Also arguing in favor of an intertask transfer approach to motor intratask analysis are some consequences emanating from attempts in verbal-learning experiments to distinguish components on the basis of simultaneously measurable physical characteristics or types of responses, as illustrated by the distinctions made between various categories of errors, i.e., omissions, intralist and extralist intrusions, etc. (e.g., Cook, 1961). While such analyses may well contribute significantly to the long-range analysis and understanding of underlying verbal-learning processes, considerable evidence has already accumulated indicating that the available component measures do not directly reflect correspondingly separate learning processes, either alone or in combination. For example, efforts to identify the learning or availability of either correct or incorrect responses on the basis of whether or not these are emitted by the subject while performing on the learning task leave much to be desired, as indicated by the considerable discrepancies between this measure and number of responses the subject is capable of producing during free response recall (e.g., McGovern, 1964). In view of this discrepancy, it also follows inevitably that attempted identification of the associative phase of paired-associate learning with the period intervening between the first occurrence of the response anywhere in the list and its correct elicitation by its stimulus term (Underwood & Postman, 1960) will similarly be doomed to failure. But what is important for present purposes is that the recognition of these noncorrespondences between measurable response properties and underlying learning processes came largely as a result of the application of intertask transfer techniques to the analysis of intratask component processes, as represented in the present illustration by the comparison between responses emitted during actual paired-associate learning and during a free response–recall test. This, of course, does not necessarily imply a similar noncorrespondence between physically measurable characteristics and behavioral processes in motor-learning tasks, although this remains a likely possibility that calls for the application of intertask transfer techniques as the only available means for empirical investigation of the question.

## Summary and Conclusions

If nothing else, the present discussion should have demonstrated research on complex intertask transfer to be fraught with apparent inconsistencies and illusory artifacts, of which the greatest must surely be the impression induced by a superficial examination of titles and topical headings that research interest in transfer is currently on the wane. Quite to the contrary, the magnitude and generality of the effects produced by previous learning upon performance on new learning tasks require that transfer phenomena be placed at or near the head of the list insofar as over-all importance to psychology is concerned. The apparent relative decline in research nominally concerned with transfer can be attributed to a number of factors, most notably its increasing overlap with investigations of single-task acquisition and retention, wherein the use of transfer concepts and techniques has become a predominant characteristic of the typical experimental methodology, especially for purposes of the analysis of component processes in complex learning tasks. A similar emphasis on component analysis also appears to be in process of development in investigations directly concerned with transfer phenomena, which have progressively shifted away from the gross measurement of overall positive or negative transfer effects toward a more precise specification and analysis of intertask component relationships which produce facilitation or interference in second-task learning.

Another promising extension of transfer techniques, which has probably received more attention here than is merited by the actual magnitude of the research effort thus far directed thereto, concerns the analysis of effects produced by intratask relationships between the various constituent elements of multiple-response learning tasks. It is already clear that such intratask transfer effects are of considerable significance even when the various elements of the task are made minimally similar to one another, and that the development of techniques for overcoming intratask interference produced by the context of other elements in the task may constitute a major part of the learning that takes place within complex multiple-response tasks. Moreover, the presence of such intratask interference is suggested by several lines of evidence to represent a significant source of intertask facilitation with respect to subsequent performance on related tasks, leading to the tentative formulation and defense of a general principle stating that *intertask facilitation is produced by intratask interference.*

Finally, the current predominance of verbal-learning studies has necessitated that the present discussion be limited primarily to research in this area. That other areas of human-learning research, such as motor-skills learning, appear not to be sharing in the recent revival of research interest evident in the verbal-learning area may well reflect the failure to accomplish a similar adaptation of transfer techniques for use in formulating and investigating fundamental problems in these other types of learning situations. Although the uses to which transfer methods and concepts have been applied effectively in verbal-learning research offer a number of suggestions as to potentially fruitful lines of investigation in motor-skills learning, some of which have been described explicitly herein, the extent to which these techniques can contribute toward the advancement of our analysis and understanding of motor-skills learning can be ascertained only through their actual use in motor-learning situations. Hopefully, one effect of the present analysis will be to stimulate future interest in the more extensive investigation and integration of transfer within other areas of learning research, similar to what is currently taking place in the area of verbal learning.

## REFERENCES

Adams, J. A. Motor skills. *Annu. Rev. Psychol.*, 1964, **15**, 181–202.

Ammons, R. B. Effects of pre-practice activities on rotary pursuit performance. *J. exp. Psychol.*, 1951, **51**, 187–191.

Annett, J., Golby, C. W., & Kay, H. The measurement of elements in an assembly task – the information output of the human motor system. *Quart. J. exp. Psychol.*, 1958, **10**, 1–11.

Archer, E. J. Effect of distribution of practice on a component skill of rotary pursuit tracking. *J. exp. Psychol.*, 1958, **56**, 427–436.

Archer, E. J. A re-evaluation of the meaningfulness of all possible cvc trigrams. *Psychol. Monogr.*, 1960, **74**, 10 (Whole No. 497).

Archer, E. J., & Bourne, L. E., Jr. Inverted-alphabet printing as a function of intertrial interval and sex. *J. exp. Psychol.*, 1956, **52**, 322–328.

Arnoult, M. D. Stimulus predifferentiation: Some generalizations and hypotheses. *Psychol. Bull.*, 1957, **54**, 339–350.

Asch, S., & Ebenholtz, S. M. The principle of associative symmetry. *Proc. Amer. Phil. Soc.*, 1962, **106**, 135–163.

Barch, A. M., & Lewis, D. The effect of task difficulty and amount of practice on proactive transfer. *J. exp. Psychol.*, 1954, **48**, 134–142.

Battig, W. F. The effect of kinesthetic, verbal, and visual cues on the acquisition of a lever-positioning skill. *J. exp. Psychol.*, 1954, **47**, 371–380.

Battig, W. F. Transfer from verbal pretraining to motor performance as a function of motor task complexity. *J. exp. Psychol.*, 1956, **51**, 371–378.

Battig, W. F. Procedural problems in paired-associate research. *Psychon. Monogr. Suppl.*, 1965, No. 1 (in press).

Battig, W. F., & Brackett, H. R. Comparison of anticipation and recall methods in paired-associate learning. *Psychol. Rep.*, 1961, **9**, 59–65.

Battig, W. F., & Brackett, H. R. Transfer from verbal-discrimination to paired-associate learning; II: Effects of intralist similarity, method, and percentage occurrence of response members. *J. exp. Psychol.*, 1963, **65**, 507–514.

Battig, W. F., & Koppenaal, R. J. Associative asymmetry in S-R vs. R-S recall of double-function lists. *Psychol. Rep.*, 1965, **16**, 287–293.

Battig, W. F., Williams, J. M., & Williams, J. G. Transfer from verbal-discrimination to paired-associate learning. *J. exp. Psychol.*, 1962, **63**, 258–268.

Battig, W. F., Brown, S. C., & Schild, M. E. Serial-position and sequential associations in serial learning. *J. exp. Psychol.*, 1964, **67**, 449–457.

Battig, W. F., Merikle, P. M., & Schild, M. E. Anagram free-recall and recognition learning, and paired-associate transfer. *J. verb. Learn. verb. Behav.*, 1965, **4**, 44–52.

Besch, N. F., & Reynolds, W. F. Associative interference in verbal paired-associate learning. *J. exp. Psychol.*, 1958, **55**, 554–558.

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.*, 1953, **49**, 229–240.

Brown, S. C. Interpair interference as a function of level of practice in paired-associate learning. *J. exp. Psychol.*, 1964, **67**, 316–323.

Brown, S. C. & Battig, W. F. Second-list paired-associate facilitation produced by addition of previously-learned first-list pairs. *J. verb. Learn. verb. Behav.*, 1966 (in press).

Bruce, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343–361.

Bugelski, B. R., & Cadwallader, T. A reappraisal of the transfer and retroaction surface. *J. exp. Psychol.*, 1956, **52**, 360–366.

Cohen, B. H. An investigation of recoding in free recall. *J. exp. Psychol.*, 1963, **65**, 368–376.

Cook, J. O. Response analysis in paired-associate learning experiments. In A. A. Lumsdaine (Ed.), *Student response in programmed instruction.* Washington, D. C.: Natl. Acad. Sci. – Natl. Res. Council (Publ. No. 943), 1961. Pp. 367–373.

Dallett, K. M. The transfer surface re-examined. *J. verb. Learn. verb. Behav.*, 1962, **1**, 91–94.

Davis, R., Wehrkamp, R., & Smith, K. U. Dimensional analysis of motion; I: Effects of laterality and movement direction. *J. appl. Psychol.*, 1951, **35**, 363–366.

Day, R. H. Relative task difficulty and transfer of training in skilled performance. *Psychol. Bull.*, 1956, **53**, 160–168.

Deese, J. *The psychology of learning.* (2nd ed.). New York: McGraw-Hill, 1958.

Ebenholtz, S. M. Position-mediated transfer between serial learning and a spatial discrimination task. *J. exp. Psychol.*, 1963, **65**, 603–608. (a)

Ebenholtz, S. M. Serial learning: Position learning and sequential associations. *J. exp. Psychol.*, 1963, **66**, 353–362. (b)

Ellis, H., & Muller, D. G. Transfer in perceptual learning following stimulus predifferentiation. *J. exp. Psychol.*, 1964, **68**, 388–395.

Fallon, D., & Battig, W. F. Role of difficulty in rote and concept learning. *J. exp. Psychol.*, 1964, **68**, 85–88.

Feldman, S., & Underwood, B. J. Stimulus recall following paired-associate learning. *J. exp. Psychol.*, 1957, **53**, 11–15.

Gagné, R. M., & Baker, K. E. Stimulus predifferentiation as a factor in transfer of training. *J. exp. Psychol.*, 1950, **40**, 439–451.

Gagné, R. M., Baker, K. E., & Foster, H. On the relation between similarity and transfer of training in the learning of discriminative motor tasks. *Psychol. Rev.*, 1950, **57**, 67–79.

Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.

Goss, A. E., & Greenfield, N. Transfer to a motor task as influenced by conditions and degree of prior discrimination training. *J. exp. Psychol.*, 1958, **55**, 258–269.

Goss, A. E., Nodine, C. F., Gregory, B. N., Taub, H. A., & Kennedy, K. E. Stimulus characteristics and percentage of occurrence of response members in paired-associates learning. *Psychol. Monogr.*, 1962, **76**, 12 (Whole No. 531).

Greenberg, R., & Underwood, B. J. Retention as a function of stage of practice. *J. exp. Psychol.*, 1950, **40**, 452–457.

Harcum, E. R. Verbal transfer of overlearned forward and backward associations. *Amer. J. Psychol.*, 1953, **66**, 622–625.

Harris, S., & Smith, K. U. Dimensional analysis of motion; V: An analytic test of psychomotor ability. *J. appl. Psychol.*, 1953, **37**, 136–141.

Holding, D. H. Transfer between difficult and easy tasks. *Brit. J. Psychol.*, 1962, **53**, 397–407.

Horton, D. L., & Kjeldergaard, P. M. An experimental analysis of associative factors in mediated generalization. *Psychol. Monogr.*, 1961, **75**, 11 (Whole No. 515).

Houston, J. P. Verbal transfer and interlist similarities. *Psychol. Rev.*, 1964, **71**, 412–414.

Jenkins, J. J. Mediated associations: Paradigms and situations. In C. N. Cofer & B. S.

Musgrave (Eds.), *Verbal behavior and learning.* New York: McGraw-Hill, 1963. Pp. 210–245.

Johnson, R. B. Recognition of nonsense shapes as a function of identification and classification training. Unpublished master's thesis, Univer. of Virginia, 1962.

Johnson, R. B. Recognition of nonsense shapes as a function of degree of congruence among components of pretraining task. Unpublished doctoral dissertation, Univer. of Virginia, 1964.

Jung, J. Transfer of training as a function of degree of first-list learning. *J. verb. Learn. verb. Behav.*, 1962, **1**, 197–199.

Jung, J. Effects of response meaningfulness (m) on transfer under two different transfer paradigms. *J. exp. Psychol.*, 1963, **65**, 377–384.

Kausler, D. N., & Kanoti, G. A. R-S learning and negative transfer effects with a mixed list. *J. exp. Psychol.*, 1963, **65**, 201–205.

Keppel, G. Problems of method in the study of short-term memory. *Psychol. Bull.*, 1965, **63**, 1–13.

Keppel, G., & Saufley, W. H., Jr. Serial position as a stimulus in serial learning. *J. verb. Learn. verb. Behav.*, 1964, **3**, 335–343.

Keppel, G., & Underwood, B. J. Retroactive inhibition of R-S associations. *J. exp. Psychol.*, 1962, **64**, 400–404.

Kimble, G. A., & Garmezy, N. *Principles of general psychology* (2nd ed.). New York: Ronald Press, 1963.

Lambert, W. E., & Jakobovits, L. A. Verbal satiation and changes in the intensity of meaning. *J. exp. Psychol.*, 1960, **60**, 376–383.

Lewis, D., McAllister, D. E., & Adams, J. A. Facilitation and interference in performance on the Modified Mashburn Apparatus. I: The effects of varying the amount of original learning. *J. exp. Psychol.*, 1951, **41**, 247–260.

Long, J. B., & Donohoe, J. W. The transfer and retroaction surface remodeled. Paper read at Midwest. Psychol. Assn., Chicago, 1962.

McAllister, D. E. The effects of various kinds of relevant verbal pretraining on subsequent motor performance. *J. exp. Psychol.*, 1953, **46**, 329–336.

McGovern, J. B. Extinction of associations in four transfer paradigms. *Psychol. Monogr.*, 1964, **78**, 16 (Whole No. 593).

Mandler, G. Transfer of training as a function of degree of response overlearning. *J. exp. Psychol.*, 1954, **47**, 411–417.

Mandler, G., & Heinemann, S. H. Effect of overlearning of a verbal response on transfer of training. *J. exp. Psychol.*, 1956, **51**, 39–46.

Martin, E. Transfer of verbal paired associates. Unpublished doctoral dissertation, Univer. of Iowa, 1963.

Merikle, P. M. The effects of stimulus and response meaningfulness in four paired-associate transfer paradigms. Unpublished master's thesis, Univer. of Virginia, 1964.

Merikle, P. M., & Battig, W. F. Transfer of training as a function of experimental paradigm and meaningfulness. *J. verb. Learn. verb. Behav.*, 1963, **2**, 485–488.

Merikle, P. M., Manatis, J. V., & Battig, W. F. The effects of degree of first-list learning and experimental procedure on transfer under three different paradigms. Paper read at East. Psychol. Assn., Atlantic City, 1965.

Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven, Conn.: Yale Univer. Press, 1941.

Morgan, C. T. *Introduction to psychology* (2nd ed.). New York: McGraw-Hill, 1961.

Murdock, B. B., Jr. 'Backward' associations in transfer and learning. *J. exp. Psychol.*, 1958, **55**, 111–114.

Noble, C. E. Meaningfulness and familiarity. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 76–119.

Osgood, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, **36**, 277–301.

Osgood, C. E. The similarity paradox in human learning: A resolution. *Psychol. Rev.*, 1949, **56**, 132–143.

Peterson, L. R. Immediate memory: Data and theory. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963, Pp. 336–353.

Porter, L. W., & Duncan, C. P. Negative transfer in verbal learning. *J. exp. Psychol.*, 1953, **46**, 61–64.

Postman, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961. Pp. 152–179.

Postman, L. Transfer of training as a function of experimental paradigm and degree of first-list learning. *J. verb. Learn. verb. Behav.*, 1962, **1**, 109–118.

Postman, L. One-trial learning. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 295–321.

Richardson, J. The relationship of stimulus similarity and number of responses. *J. exp. Psychol.*, 1958, **56**, 478–484.

Robinson, E. S. The 'similarity' factor in retroaction. *Amer. J. Psychol.*, 1957, **39**, 297-312.

Rubin, G., & Smith, K. U. Learning and integration of component movements in a pattern of motion. *J. exp. Psychol.*, 1952, **44**, 301–305.

Sidman, M. *Tactics of scientific research*. New York: Basic Books, 1960.

Thompson, A. S. The effect of relevant and irrelevant familiarization on recall and recognition of responses in paired-associate learning. Unpublished master's thesis, Univer. of Virginia, 1965.

Twedt, H. M., & Underwood, B. J. Mixed vs. unmixed lists in transfer studies. *J. exp. Psychol.*, 1959, **58**, 111–116.

Underwood, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49–60.

Underwood, B. J. An evaluation of the Gibson theory of verbal learning. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961. Pp. 197–217.

Underwood, B. J. Stimulus selection in verbal learning. In C. N. Cofer & B. S. Musgrave (Eds.), *Verbal behavior and learning*. New York: McGraw-Hill, 1963. Pp. 33–48.

Underwood, B. J. Degree of learning and the measurement of forgetting. *J. verb. Learn. verb. Behav.*, 1964, **3**, 112–129.

Underwood, B. J., & Postman, L. Extraexperimental sources of interference in forgetting. *Psychol. Rev.*, 1960, **67**, 73–95.

Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning*. Philadelphia, Pa.: Lippincott, 1960.

Underwood, B. J., & Schulz, R. W. Studies of distributed practice; XXI: Effect of interference from language habits. *J. exp. Psychol.*, 1961, **62**, 571–575.

Underwood, B. J., Runquist, W. N., & Schulz, R. W. Response learning in paired-associate lists as a function of intralist similarity. *J. exp. Psychol.*, 1959, **58**, 70–78.

Vanderplas, J. M., Sanderson, W. A., & Vanderplas, J. N. Some task-related determinants of transfer in perceptual learning. *Percept. mot. Skills*, 1964, **18**, 71–80.

Wimer, R. Osgood's transfer surface: Extension and test. *J. verb. Learn. verb. Behav.*, 1964, 3, 274–279.

Winnick, W. A., & Dornbush, R. L. Role of positional cues in serial rote learning. *J. exp. Psychol.*, 1963, **66**, 419–421.

Yelen, D. R., & Schulz, R. W. Verbal satiation? *J. verb. Learn. verb. Behav.*, 1963, 1, 372–377.

Young, R. K. A comparison of two methods of learning serial associations. *Amer. J. Psychol.*, 1959, **72**, 554–559.

Young, R. K. The stimulus in serial verbal learning. *Amer. J. Psychol.*, 1961, **74**, 517–528.

Young, R. K. Tests of three hypotheses about the effective stimulus in serial learning. *J. exp. Psychol.*, 1962, **63**, 307–313.

Young, R. K., & Underwood, B. J. Transfer in verbal materials with dissimilar stimuli and response similarity varied. *J. exp. Psychol.*, 1954, 47, 153–159.

# Facilitation and Interference

## Comments on Professor Battig's Paper

PAUL W. FOX
*University of Minnesota*

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B. C. Road

Battig has treated the development and contemporary use of transfer methodologies in a scholarly and systematic manner. The dominant theme of his paper, as I see it, is that analysis of component processes within a task represents a highly fruitful approach to the understanding of between-task phenomena. This theme has been illustrated by the role that component process analysis is playing in explicating interference and facilitation effects in verbal learning situations. Two implications for motor-skills research have been drawn. First, interference in the performance of a second task may be obscured or lessened by sources of interference within the first task. This possibility is explicitly stated by Battig and follows directly from his hypothesis that intratask interference leads to intertask facilitation. The second implication is less specific. It concerns the observation that in verbal research introduction of a second task has been an effective means of analyzing learning processes going on in the first. The suggestion is made that similar procedures may be usefully applied to motor-skills research.

Both implications assume continuity rather than discontinuity between motor and verbal learning. In this respect, the failure to produce substantial negative transfer in skills learning clashes with the successful generation of interference in verbal paired-associate paradigms. With high S-R meaningfulness, for example, the Merikle data described by Battig display over 100% negative transfer under A-C and $A\text{-}B_r$ test conditions. Battig has approached this apparent verbal-motor discontinuity by way of the intratask interference–intertask facilitation hypothesis. That is, the stimulus and response relationships of devices such as star discrimeters and complex coordinators supposedly produce powerful sources of intratask interference. Thus performance on a second task is facilitated and interference is masked even when new responses are required to old stimuli. It should be noted, however, that findings of transient or nonexistent interference emerge from a considerable diversity of laboratory devices and experimental designs.[1] For example, failure to show persistent intertask decrements has characterized investigations using tasks with one, two, or three response components. Complete and partial response changes in all possible combinations have been introduced in both retroactive and proactive designs. Increased interpolated training on altered response elements produces immediate postshift decrements during relearning, but the losses rarely fall to initial original learning levels and are quickly overcome. Increasing amounts of original training facilitates performance on various second-task partial response reversals. Barch and Lewis (1954) show joint facilitation (in time on target) and interference (persistence of errors) when subjects are shifted from one task to another demanding different responses. This finding is consistent with other reports from the research program of Lewis and his associates at Iowa. In a more recent study designed to maximize the probability of producing interference, Blick and Bilodeau (1963) controlled the magnitude and direction of interpolated responses according to quantitative similarity variations modeled on Osgood's transfer and retroaction surface. Similar to prior efforts to generate interference with uncontrolled response interpolations, these authors report null results.

Battig's hypothesis of intratask interference leading to intertask facilitation must be stretched precariously thin to cover the above results. It assumes powerful sources of intratask interference operating in tasks ranging from simple line drawing to complex coordination. Further, the hypothesis suggests that such sources of interference exist under the variety of experimental conditions noted above and in all cases produce a masking of possible intertask interference. At the risk of placing my foot (motor element) in my mouth (verbal element), I wish to approach this problem in a somewhat different way. Namely, I will try to show that little discontinuity exists between motor and verbal findings, and that in fact the data are complementary rather than conflicting. The viewpoint proposed is that empirical differences between motor and verbal transfer studies may reflect (a) possible procedural differences in the application of specific S-R relationships; and (b) characteristic differences in the selection and expression of transfer indices.

The first possibility (i.e., noncorrespondent motor-verbal application of interference producing S-R relationships) would receive some support if it could be shown that new stimuli as well as new responses were characteristic of second-task response alterations. To rephrase in symbols borrowed from verbal learning: To what extent have the motor manipulations described above more closely conformed to second-task C-D than to A-C relationships? In terms of nominal stimuli, this suggestion is of course patent nonsense. As has been noted elsewhere, however (Bilodeau, 1966), the critical stimulus may well be some function of response error (i.e., information feedback) rather than the nominally displayed target indicators. This suggestion is also in accord with earlier explanations of transient interference based upon correlational analyses. Namely, under various complete or partial response changes the subject is learning skills common to all tasks (Bilodeau & Bilodeau, 1961). Thus, in terms of functional S-R relationships (Underwood, 1963), subjects may at worst be performing under second task C-D conditions and at best under more facilitating relationships.

---

[1] The summary sentences following refer generally to experiments performed by Professor Don Lewis and his associates at Iowa. Specific citations have been omitted since much of this research is referenced by Bilodeau & Bilodeau in their 1961 review and by Battig in the preceding paper.

In line with this viewpoint, motor response tasks closely modeled after verbal paradigm procedures should produce complementary findings. Data from an experiment performed by Duncan (1958) illustrate a family of such tasks, and also bear upon the hypothesis proposed by Battig. Subjects learned a paired-associate task with 13 similar forms as stimuli and a simple, discrete, stick movement into one of 13 slots as responses. One condition received 200 trials, 20 per day, with the same form-movement pairings (Constant task treatment). A trial was defined as one run through the 13 pairs. Other subjects received 200 trials, but on each successive day for 10 days a new set of responses was required to the original 13 stimuli (Re-paired treatment). Essentially, this was a variant of a repeated A-B$_r$ situation. Finally, other subjects were trained with from 2 to 10 different sets of completely new stimuli and movement responses (Varied treatment). Test trials on days 11 and 12 called for acquisition of movement responses to forms and cvc's entirely dissimilar to prior training stimuli. Findings relevant to the present discussion were (a) highly significant improvement for all conditions from day 1 to 10 (even for successive S-R scrambling in the Re-paired treatment); (b) positive transfer on the test trials, in comparison with day 1 training performance, and more so for Varied than Constant or Re-paired treatments; and (c) test trial facilitation of approximately equal amounts for the Constant and Re-paired task subjects. Duncan hypothesized that training experience with many different sets of stimuli developed instrumental observing responses important to the Varied treatment's transfer superiority. He concluded that amount of facilitation "... in the simple sense of intertask improvement, ... should vary directly with degree of intralist, and perhaps interlist, similarity, and inversely with meaningfulness of stimuli" (Duncan, 1958, p. 72). Although derived from differing situations, Battig's principle appears to receive more than superficial support from Duncan's data, and their independent hypotheses emphasize applicability to a similar set of conditions and learning processes.

It was noted earlier that differing indices and expressions of transfer may obscure as well as reflect similar findings. Thus, at least some of the apparent nconsistencies between and within verbal and skills areas may be related to his factor. Because of its accessibility, I will try to illustrate this point by reference to Battig's Table I presentation of data supplied by Merikle. The formula used in computation of percentage transfer in Table I is shown below along with data from the first column of the Table (A-C paradigm) illustrating ts application, (C stands for control, T for the experimental treatment, and, in the case of error data, total possible has been entered as zero).

$$\frac{\text{Control} - \text{Treatment}}{\text{C}-\text{Total possible}} = \frac{5.00 - 13.13}{5.00 - 0.00} = \frac{-8.13}{5.00}(100) = -162.6\%$$

Expression of total possible errors as zero essentially converts the original formula (No. 6 in Gagné, Foster, & Crowley, 1948) to their formula No. 2, the atter being influenced by the absolute value of the scores entered. In view of this, what change in the numerator is most appropriate to any given change in the denominator? Is the prediction made that the A-C treatment will make more errors than the C-D control, five more errors than the control, or more than twice as many errors as the control? If the C-D control were to have made 10 errors, an event seemingly possible with any number of procedural changes, should one predict that the A-C condition will make 18.13 errors or 26.26 errors? As must by now be painfully obvious, −162.6 percent transfer can, in such a case, be obtained only by *doubling* A-C errors. An increase of only five errors for both treatments reduces the obtained transfer to −82 percent, or half its former magnitude. The present writer is familiar with no case, however, in which differential transfer predictions are predicated upon ratio relationships.

This example has been treated in detail in order to illustrate a more general issue. The decision to express intertask interference or facilitation effects in terms of second task origin (immediate transfer), slope (delayed transfer), asymptote (restricted in tasks with limited ceilings), trials saved, or some ratio of gain or loss may dramatically influence the magnitude of one's findings. In the absence of clearly defined and standard indices of transfer, comparison between studies both within and between areas is severely handicapped. One final note (again in reference to the Table I data), by way of concluding this point and introducing the next. High-meaningfulness S-R subjects made an average of 13.68 errors during the learning of List 1. In terms of savings, a measure often applied in motor transfer research, all paradigms thus show List 2 benefits. In verbal research this index of transfer is seldom applied since the C-D reference condition provides a control for nonspecific facilitation effects. Since a previous section of this discussion likened certain interference designs in motor learning to just this C-D condition, the intertask effects involved require further elaboration.

Belief in the facilitating effects of previous laboratory experience upon present and even dissimilar tasks is widespread, as Battig has noted. The ubiquitous nature of such nonspecific transfer is attested to by consideration of practice effects in the method section of standard psychological reports. The present remarks will try to show that such facilitation (a) represents one of the most reliable and substantial of intertask effects; and, (b) is among the least studied and understood phenomenon in this field. If such an appraisal is accurate, nonspecific intertask facilitation is a major source of difficulty for interference-producing manipulations in motor-skills research. The relative unpopularity of research on nonspecific transfer may in part be due to its apparent correspondence with older and less analytic approaches. The traditional question of whether practice on task A will facilitate performance on task B is not, by contemporary standards, very informative. Fortunately, several recent investigations have been initiated solely to identify the specific component processes in learning to learn. Postman and Schwartz (1964) trained subjects on serial or paired-associate tasks with trigram or adjective materials. Intra- and intertask item similarity was reduced as much as possible, and second-task training called for serial or PA adjective learning. In comparison with first-task performance, all treatments showed second-task facilitation. Moving from one task to another produced less facilitation than staying with the same task, whether serial or PA. The change in materials produced less effect, though no subjects were switched from adjectives to trigrams. Second-task performance differences showed up later than earlier, and the authors conclude that differences during the associative stage are primarily responsible for the increased facilitation. No subjects were tested on trigrams and this conclusion may need elaboration. It does reflect, however, a more analytic approach to the general phenomenon.

Postman (1964), in the second of the learning-to-learn series, examined PA adjective performance across three successive sets of first and second lists under A-B′, A-C, A-B$_r$, and C-D paradigms, respectively. First-list learning for each set was on A-B, where the first letter designates the stimulus, the second letter the response, and B′ an adjective associatively related to B. The findings are complicated, but several of them appear particularly relevant to the present discussion. First, speed of first-list learning increased at comparable rates for all paradigms. Second, the C-D reference treatment displays somewhat greater facilitation over the first two test sets and a steeper acquisition slope on the third than the other treatments, even A-B′. That is, whatever factors produce improvement in C-D, they are often more effective than the deliberate S-R manipulations chosen to facilitate treatment A-B′. Third, and finally, this experiment allows examination of initial shift and total test performance, and also delayed transfer effects. It is the latter which are, in this case, most informative. Namely, while the initial test trial of A-C performance is far below that of A-B′, A-C (interference paradigm) subjects show delayed *positive* transfer in relation to treatment A-B′ (a facilitation paradigm).

These studies serve to illustrate several points. First, prior practice represents a powerful source of intertask facilitation, even under conditions of specifically designed interference. Second, post-shift decrements, as in the A-C condition, may quickly convert to positive delayed transfer. Both findings, it may be recalled, were described earlier as characteristic of motor transfer. These results may be subsumed under the heading of intertask facilitation. When emphasis is directed toward retention rather than learning, however, a curious inconsistency arises. Battig has acknowledged this by urging the methodological distinction between transfer and proaction procedures. The extensive nature of this particular problem merits further discussion.

Intertask proactive facilitation is generally viewed as displaying positive slope and negative acceleration when degree of facilitation is plotted over practice on successive tasks. The gain may be completely realized on the second task (Greenberg & Underwood, 1950) or may still be showing up late in the series (Duncan, 1960). In contrast, repeated experience on nominally dissimilar tasks may lead to marked decrements or proactive interference if performance is measured in terms of retention. This has been shown by Greenberg and Underwood (1950), and convincingly reviewed by Underwood (1957). The phenomenon has also been obtained in studies of short-term memory as a function of number of preceding items, and has been discussed most recently by Keppel (1965). At least a minor paradox seems involved. That is, successive practice leads to intertask facilitation in the learning of new material, and to intratask interference in the retention of old material. It may be, of course, that, by increasing the amount of practice on successive lists (2, 3, . . ., $n$) to that amount required for originally learning list 1 (with concomitant successive degrees of overlearning), retention of any one list after 24 hours will show no such decrement. In at least one such experiment, however, this was clearly not the case and other explanations are required.[2] Since Battig has illustrated his provocative hypothesis with data from a recognition test (retention) and suggested its applicability to second-task acquisition phenomena (learning), the issue involved is seen to turn up in unexpected ways. In short, it is not enough to distinguish methodologically between transfer and proaction designs. A need also exists for further investigation aimed at unraveling the effects of proactive facilitation in acquisition on the one hand and proactive interference in retention on the other.

---

[2] The cooperation of Dr. Jenkins and Miss Brown in bringing these unpublished data to the writer's attention is gratefully acknowledged.

The preceding material has been concerned with two rather general problems. The first, represented by the apparent empirical differences obtained with motor-skills and verbal paradigm researches, was explicitly noted. The conflict has been approached by way of differences in procedures and in measures of transfer. Battig's hypothesis was discussed in this context, and tangential evidence provided for the view that it has relevance to motor studies closely modeled on verbal paired-associate procedures. Nonspecific transfer was discussed as a powerful intertask source of facilitation in both task areas. The argument has been that obtained differences are more superficial than basic. The other problem concerned intertask processes which appear both reliable and substantial and which seem to require increased research efforts. Learning to learn and the paradoxical effects of prior experience on learning and retention fall within this category.

The second implication for motor-skills research drawn by Battig from his coverage of current verbal approaches concerns the use of a second task in order to investigate component learning processes in the first. There is some evidence to support the notion that a majority of motor transfer experiments have been designed with similar objectives in mind. Fitts (1964), for example, has described early, middle, and late acquisition stages in motor-skills learning in terms derived from verbal component process analyses. Learning stage analysis is of relatively recent vintage even in verbal learning, and identical terminologies and approaches are rare in motor skills. However, recent motor learning reviews (Adams, 1964; Bilodeau & Bilodeau, 1961) clearly emphasize the learning-performance distinction rather than second-task positive or negative transfer effects. Another way of saying this is that such experiments are mainly concerned with delineating learning as opposed to performance variables. Transfer designs play a key role in such analyses. I suspect, further, that the increasingly widespread use of transfer methodologies noted by Battig reflects at least part of this reorientation. Such speculation appears to be more than casually justified by the past decade of skills research.

In concluding this discussion, I wish to describe several promising research techniques for studying interference and facilitation not previously touched upon by Battig. The techniques to be discussed involve S-R manipulations producing interference or facilitation not between traditional paradigms, or even within any given paradigm with variations in meaningfulness. Instead, they emphasize S-R relationships within paradigms and within subjects when meaningfulness is constant. Their orientation is based not upon manipulation of two points on a similarity dimension (identical and neutral), but upon associative relations between words scaled along a probability continuum from 1.00 to zero.

Bilodeau has already described the general methodology of stimulated verbal recall (pp. 337–345) and the procedures used by Blick (1965) to produce unusually large amounts of interlist interference. This effect, obtained in an ABB design, occurred as a function of the pre-experimental free associative strength of the primary response ($R_1$) in conjunction with experimental exposure of the primaries on List 1 during training. Three subsequent experiments (Fox & Bilodeau, 1966) have shown that Blick's procedure may be used to generate interlist *facilitation* rather than interference when the $R_1$ words have associative connections with the List 2 words to be recalled. That is, when a subject is trained with $R_1$ words on one list, and secondary associates or some other words on a second list (e.g., $R_c$), the probability of subsequent correct retrieval of the $R_c$ words is a joint function of $p(R_1)$ and $p(R_1 \rightarrow R_c)$. If the first term is high and the second low (as in Blick), correct retrieval is depressed and many $R_1$ words intrude. If the first *and* second terms are high, production of $R_c$ in recall is facilitated and far fewer $R_1$ words intrude. The critical point is that the same paradigm and operations may lead to either interlist interference or interlist facilitation, dependent upon the internal associative probability connections between terms.

A second line of research has investigated proaction and retroaction effects when the materials employed bear known strength in the individual's own response hierarchy. Coleman (1963) assessed individual associative hierarchies through measures of response latency, and used the materials so derived in a subsequent paired-associate task. He found that an overwhelming majority of intrusion errors during learning could be specified in terms of the individual subject's pre-experimental response hierarchy. Paired-associate learning of materials with low extraexperimental interference, as previously defined, was more rapid than paired-associate learning of material with known sources of high extraexperimental interference.

In general, experiments relating intra- or interlist verbal recall to word association responses typically rely upon assessments of pre-experimental normative data. Coleman's approach is an exception, and studies using the same subjects in both norm gathering and recall phases are rare. This writer, in an investigation stimulated by E. A. Bilodeau, interpolated the individual subject's own free-associative primary response between training and recall phases (Fox, 1964). The salient results showed within-subject facilitation ($p$ recall $= .84$) and interference ($p$ recall $= .26$) as functions of the relationship between training and interpolated material, and the free or controlled regulation of the recall environment.

A third line of current activity is illustrated by Brown, Jenkins, and Lavik, (1965). They use interspersed test item techniques (modeled after stimulus generalization procedures) following paired-associate training. The results display decreasing generalization gradients of response transfer as a function of the decreasing associative relationship between original and interspersed test stimuli. All in all, these examples display an increasing concern with pre-experimental and experimental assessment of associative word relations. Viewed in conjunction with the techniques described by Battig, they show the increasingly wide range of analytic procedures now available for specifying the variables and conditions governing interference and facilitation.

## REFERENCES

Adams, J. A. Motor skills. *Annu. Rev. Psychol.*, 1964, **15**, 181–202.

Barch, A. M., & Lewis, D. The effect of task difficulty and amount of practice on proactive transfer. *J. exp. Psychol.*, 1954, **48**, 134–142.

Bilodeau, E. A. Acquisition of skill. In *International encuclopedia of the social sciences.* 1966 (in press).

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Blick, K. A. Cultural primaries as a source of interference in short-term verbal retention. *J. exp. Psychol.*, 1965, **69**, 246–250.

Blick, K. A., & Bilodeau, E. A. Interpolated activity and the learning of a simple skill. *J. exp. Psychol.*, 1963, **65**, 515–519.

Brown, L. K., Jenkins, J. J., & Lavik, J. Response transfer as a function of verbal association strength. *J. exp. Psychol.*, 1966 (in press).

Coleman, E. B. The association hierarchy as an indicator of extraexperimental interference. *J. verb. Learn. verb. Behav.*, 1963, **2**, 417–421.

Duncan, C. P. Transfer after training with single versus multiple tasks. *J. exp. Psychol.*, 1958, **55**, 63–72.

Duncan, C. P. Description of learning to learn in human subjects. *Amer. J. Psychol.*, 1960, **73**, 108–114.

Fitts, P. M. Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 244–286.

Fox, P. W. Stimulated recall as a function of interpolated word associations. Unpublished doctoral dissertation, Tulane Univer., 1964.

Fox, P. W., & Bilodeau, E. A. Associative rules governing recall and misrecall. *J. exp. Psychol.*, 1966 (in press).

Gagné, R. M., Foster, H., & Crowley, M. E. The measurement of transfer of training. *Psychol. Bull.*, 1948, **45**, 97–130.

Greenberg, R., & Underwood, B. J. Retention as a function of stage of practice. *J. exp. Psychol.*, 1950, **40**, 452–457.

Keppel, G. Problems of method in the study of short-term memory. *Psychol. Bull.*, 1965, **63**, 1–13.

Postman, L. Studies of learning to learn, II: changes in transfer as a function of practice. *J. verb. Learn. verb. Behav.*, 1964, **3**, 437–447.

Postman, L., & Schwartz, M. Studies of learning to learn, I: transfer as a function of method of practice and class of verbal materials. *J. verb. Learn. verb. Behav.*, 1964, **3**, 37–49.

Underwood, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49–60.

Underwood, B. J. Stimulus selection in verbal learning. In C. N. Cofer and B. S. Musgrave (Eds.), *Verbal behavior and learning.* New York: McGraw-Hill, 1963. Pp. 33–48.

# Chapter 6

# Information Feedback

INA McD. BILODEAU
*Tulane University*

## Labels and Definitions

A variety of general terms has been used for much the same experimental procedures: *knowledge of results* (Dees & Grindley, 1951; Lavery & Suddon, 1962); *feedback* [achievement information feedback (Smode, 1958); information feedback (Bourne, 1957); reinforcing feedback (A. Taylor & Noble, 1962); psychological feedback (Payne & Hauty, 1955); etc.]; *reinforcement* (Suppes & Frankmann, 1961; Taubman, 1944); and *reward* (C. E. Noble & Alcock, 1958; Saltzman, 1951). An experimenter's choice among these terms sometimes suggests the essential function he attributes to feedback (is determined by his systematic position), but, as often, his choice of a given term seems to imply little about his systematic commitment. There is some tendency for *feedback* to be used in tracking studies, and *knowledge of results* (KR), *information feedback* (IF), or *reinforcement* to be used in studies of positioning and selective learning, but examples of the opposite can also be found. Different terms may be used very similarly by different investigators, and the same term may have different meanings to its different users. For example, H. Seashore and Bavelas (1941) apparently approve *knowledge of results* because it carries some connotation of the subject's inner workings; E. A. Bilodeau (1956) has used the same term (abbreviated to KR), but clearly wishes to strip it of the very connotation of which Seashore and Bavelas approve and to restrict the definition of KR to external manipulations. Like Bourne and his associates (1963), who use *informative feedback* (IF), Bilodeau puts the subject's perceptions, reasonings, and interpretations among the dependent events, to be studied as consequences of stimulus (S) events (whether called IF or KR). Beside general terms, there are a number of specific terms in reasonably wide use, usually produced by adding a modifier to one of the basic terms. The modifier often merely denotes a specific manipulation, as for example, *supplementary*, *intermittent*, *augmented*, and *terminal* feedback; but sometimes the specific term has systematic meaning as well, as, for example, in uses of *intrinsic* vs. *extrinsic* and *action* vs. *learning* feedback. The operationally aimed specific terms will be treated in later sections; a discussion of general terms, definition, and specific terms that make a systematic commitment continues below.

Choosing a term is arbitrary and feedback's functions in learning and performance are matters of dispute, but there seems no excuse in contemporary psychology for confusing external stimuli with the subject's responses. Using the same term for the experimental processes of providing knowledge of results (Brown, 1949), i.e., making information available, and for the recipient's processing the information is conducive to misunderstanding and discord. Obviously, if *knowledge of results* refers to the subject's awareness of the effects of his responses, some other term is needed for the variables that in turn control the subject's awarenesses and evaluations. In this sense, *knowledge of results* is a construct, with the usual position between antecedent, or independent, and consequent, or dependent, events, and the antecedent events and the construct need different names. The excuse for the confusion seems to be that *knowledge of results* is a very bad and misleading expression and the conclusion is that the dangerous expression should be discarded. Unfortunately, KR can be discarded only in favor of less offensive expressions, and not for an absolutely favorable term. *Reinforcement* irritates some workers, who associate the word with theoretical positions on how the reinforcement acts (E. A. Bilodeau & Bilodeau, 1958a). *Feedback* unfortunately has clouds of servomechanism wreathing its head, and is (often with sensory or sensorimotor added) a legitimate and useful expression for response-produced stimulation which to the present writer is part of the response (see Frequency of IF, below). Perhaps *IF* is the least noxious of the available expressions, if only because *information* has a more contemporary ring than *knowledge*.

As we shall use the term (and with no originality), *IF* refers to stimuli under the experimenter's control, and related to response (R), presented during the course or at the end of a response. The subject's cognizance of what he is doing or has done is not IF, nor is the proprioceptive stimulation produced by the behavior, though one may use proprioceptive IF (Lincoln, 1956). The discrepancy between goal and obtained responses is the IF (Fitts, Noble, Bahrick, & Briggs, 1959; E. A. Bilodeau, 1953b; Suppes & Frankmann, 1961; Rouanet & Rosenberg, 1964), and varying the relationship of the IF scale to the scale of the experimenter's records is the most important area of research in IF; that is,

$$\mathrm{IF}_n = f(R_n - G), \tag{1}$$

where $R_n$ is a quantitative expression of the response on Trial $n$ and $G$ is the goal response, or value, scored as errorless (a hit, coincidence of manipulandum and target, etc.). The $f$ may convert the IF to um-hm vs. uh-uh, at one extreme, and be the complex of display features in continuous tracking at another. The object of the experimental work is to determine the relationship between IF variables and behavior, i.e., to express

$$R_{n+1} = f(\mathrm{IF}_n) = f(R_n - G). \tag{2}$$

## Hypothetical Properties of Feedback

Probably most psychologists would allow that IF has at least the following three empirical properties, regardless of hypothesized theoretical properties: (a) R strengthening, (b) sustaining performance, and (c) eliminating previously established Rs. As for its theoretical properties, logically, IF, as any stimulus, can have all or any of three: (a) directive, (b) motivating, and (c) reinforcing. A glance through the discussion sections of research done in the 1930's (for example, Gilbert & Crafts, 1935; Trowbridge & Cason, 1932), will quickly establish both the age and venerability of the hypothetical roles of IF in learning and performance, and the conviction that little forward progress has been made in relating IF exclusively to stimulus, drive, or reinforcement in the past 30 years. The really fortunate investigators are those who began to study feedback from an interest in probability learning, and have felt free to plunge at once into studying response probability and response variability as functions of a variety of feedback manipulations.

Brown has made the clearest presentation of the theoretical status of feedback and its alternative hypothetical properties (1949), and his thorough discussion of problems and variables deserves more attention than its obscure publication has permitted. Among other investigators, the directive property seems generally accepted, though not with equal stress by all. E. A. Bilodeau, who emphasizes the directive property, maintains one extreme position. Others (Adams, 1964; C. E. Noble & Broussard, 1955) are more moderate, either allowing all three properties as reasonable or not committing themselves as to how IF operates. Withholding commitment until more low-order laws are gathered is doubtless the sensible position; few manipulations have yet been offered that would vary the potential properties separately, and the usual prediction seems to be resolutely independent of the property, accounting for it. Taken as a stimulus that provides a quantified error score, IF is at once the stimulus to which the subject is to respond (his cue) and his reinforcement. Motivationally, IF can operate either by being there or by being variable. Conceding that IF clearly has empirical cue properties, one might still push the question of whether the cuing depends upon an incentive or reinforcing IF. If one turns to the data to settle IF's theoretical action, the situation is as bad. Reducing, eliminating, or impoverishing IF from some source lowers performance, because the S is not there to evoke R, the source of drive has been removed, or the reinforcing agent has been eliminated. If, despite decrement, performance is better than it would have been without the period of richer IF, have the subjects learned more because the S to evoke the relevant R was once present, they were more motivated during acquisition and therefore put out more responses, or they had more frequent reinforcement in the past? (A food pellet can have the same three properties and pose as much of an explanatory problem as + 10°–i.e., the people studying IF are not alone in having trouble with their hypothetical properties.) A relatively small group of studies avoids much of this confounding by using a two-choice situation and "right" vs. "wrong" as IF (J. T. Spence, 1964), or IF for correct vs. incorrect Rs (Chinn & Alluisi, 1964). The information in the IFs is the same, and yet the behavior to the two is not.

Perhaps part of the difficulty in separating IF's motivating and cue properties is that attempts to demonstrate motive-incentive effects from instructions, tuitions, shocks, praise, exhortations, and extraordinary IFs usually fail. It is hardly surprising, then, that dramatic motivational effects are not readily evidenced after the dry report of an error by size and direction. A few demonstrations (whose theoretical implications might be wrangled over) have succeeded, of course. Threatening the subject with a long or arduous practice session usually slows him down (Ross & Bricker, 1951; Saufley & Bilodeau, 1963). Fleishman (1958) combined an arduous-practice threat with various other incentive techniques and produced slightly higher performance levels, but only in his high-ability subjects. C. E. Noble, Fuchs, Robel, and Chambers (1958) succeeded in producing social competition effects on a self-paced psychomotor test, not on an experimenter-paced test, and have a good discussion of possible reasons why. French (1955) has found that incentive instructions can work when they are suited to the subject's motivations. Payne and Hauty (1955) used incentive-IF techniques to maintain high levels of output over hours of compensatory tracking; their success in producing such an effect may be just because they did use a day's, rather than a few minutes', practice. Long-term studies, at any rate, seem more promising than studies in which one group is told for 10 minutes that it is doing very well and another group is told that it is doing very poorly. An example of the rewards of long-term effort is Walton and Begg's (1958) 20-day study of imbeciles on a fatiguing task; these investigators were able to observe differences among encouraged, goal, and competing groups in the first half of practice, and a later decline.

Despite previously expressed objections (E. A. Bilodeau & Bilodeau, 1958a, 1961), the writer now admits the harmlessness of including IF in the general class of reinforcing events, without commitment on its theoretical action – i.e., IFs are certainly stimuli that are consequences of behavior and that serve to modify R probability. While the utility or advantage gained in using *reinforcement* rather than *IF* may be questioned, there seems to be no impropriety in substituting one empirical term for another. The several meanings of *reinforcement*, however, and the argument over them, are enough to frighten the writer away; when *reinforcement*$_1$, *reinforcement*$_2$,..., and *reinforcement*$_n$ are disentangled seems adequate time to attempt a translation.

## An Area Delimited

The topic of IF, or the stimulus consequences of behavior, is so broad that it puts no limits to one's coverage. The writer therefore chooses to emphasize IF and positioning responses in discrete-trial presentation. Heavier and lighter behavior are treated only incidentally. Among the rationalizations for this arbitrary restriction, the most presentable is avoiding redundancy: Noble's *Selective Learning* (Chapter 2), Poulton's *Tracking Behavior* (Chapter 8), Smith's *Behavioral Cybernetics* (Chapter 9), all have specific feedback contributions; and L. Bourne, discussant for this chapter, fills in some important mathematical learning and concept-formation holes. The following presentation is drawn largely from research on motor responses that can vary in extent or degree along some continuous dimension, can be scored along the dimension for the experimenter's record, and a function of the score provided to the subject, e.g., the subject sets a control; a single setting is made per trial; some noticeable break occurs after one response, and before the next is called for; and ordinarily the IF occurs in the break between responses. Figure 1 presents schematically the sequence of events in a typical positioning task with discrete trial-presentation.

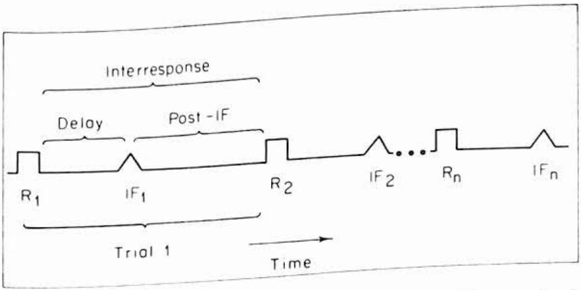


FIG. 1. Typical sequence of events in a positioning task.

Four major areas of research are reviewed: *locus* of IF—its temporal placement with respect to the R to which it refers and to the next R in the sequence; *frequency* and *frequency schedules* of IF—absolute and relative numbers of IFs and presence vs. absence of IF; *augmented IF*—adding IF to the standard IF the task provides; and *IF transformations*—manipulations of the function relating IF value to R.

In most of the contemporary studies reviewed below, precautions have been taken to minimize transfer of training from past experience by such techniques as telling the subject he is to learn to draw a line of the *correct* length or move a lever the *correct* distance (not to learn to draw a 3-inch line or make a 60° pull); and expressing goals, Rs, and errors in arbitrary units, rather than in inches, degrees of arc, or other familiar units. The subject also is ordinarily blindfolded (or fixation-point and screens are arranged to prevent his watching any effect of his R during its execution)—unless action-IF treatments are deliberately introduced for comparison; like controls are in effect for unwanted auditory, tactile, and other cues that the apparatus might contribute. Two procedures are typical. If error is to be reported, the subject may be, for example, advised to try to make an R of the correct amplitude or to try for an R of 45 *units* (which might be 67.5°) and told that he is 5 *units* (not 7.5°) above or 4 *units* (not 6°) below the goal. If R amplitude is to be reported, the subject may be told to try for a score of 45 units (not 67.5°) and IF is 50 units (not 75°), 41 units, etc. (The word *unit* may itself, of course, be omitted.) Experimenter, control, and IF scale are all concealed from the responding subject in the standard terminal-IF treatment, i.e., research on IF with positioning Rs is concerned with *operational* feedback (Smith, Chapter 9, p. 465).

## Temporal Locus

The locus of IF refers to its position in the trial cycle, and includes delay of IF. Except perhaps for feedback's role in learning and the related question of the function of supplementary IF, no other variable has aroused as much dispute as delay—the time elapsing between a response and its IF. Some of the controversy comes from the data, much from the implicit or explicit expectation that temporally displacing IF from the response would produce a gradient similar to gradients obtained by temporally displacing primary reinforcement from an instrumentally conditioned response. With a ready, though sadly misapplied, theoretical basis, delay was the most interesting, even the only really interesting, feedback variable of the 1940's. A closer look at the theoretical sections of reports of gradients in instrumental conditioning would have revealed that much of the point was to identify variables that controlled slope and extent of gradients of delay (Grice, 1942); and in contexts of both feedback (Lorge & Thorndike, 1935) and food delay (K. W. Spence, 1956), the efficacy of empty delay, or *delay alone*, has been doubted. Studies of simple positioning responses support the doubts of the importance of sheer delay of IF – to the conclusion that the most widely interesting manipulation is among the least effective. Because of its thus hallowed status as an area of dispute, and its age, temporal locus is treated as a separate topic, rather than as an item under the more general topic of transformation; K. U. Smith, Wargo, Jones, and Smith (1963), however, quite properly treat temporal and spatial displacements together.

### DELAYS IN TRACKING

Effective delays of feedback are not difficult to find. In tracking (Warrick, 1949; Conklin, 1957; Garvey, Sweeney, & Birmingham, 1958), in handwriting (van Bergeijk & David, 1959; W. M. Smith, McCrary, & Smith, 1960), in speech (K. U. Smith, 1962), in selective learning (C. E. Noble & Noble, 1958), and, generally, wherever the subject responds continuously to a continuously changing display (of which his feedback is a part) or receives terminal, lump-sum IF after a series of Rs, temporal lag in the response-feedback sequence disrupts behavior. (If a subject gets a score only after he positions a stick, R error is a suitable index of performance; but if the subject receives IF as he positions the same control, response time must be scored if improvement with practice is to be demonstrated.) To illustrate the disruptive effects delayed IF can have, Fig. 2 presents an example of serial and terminal feedback in selective learning (C. E. Noble & Noble, 1958). Otherwise, the reader is referred to Chapters 8 and 9 for more extensive review of effective temporal lag in feedback and the section on Augmented IF below for further elaboration of continuous vs. terminal treatments. The writer has, however, de-

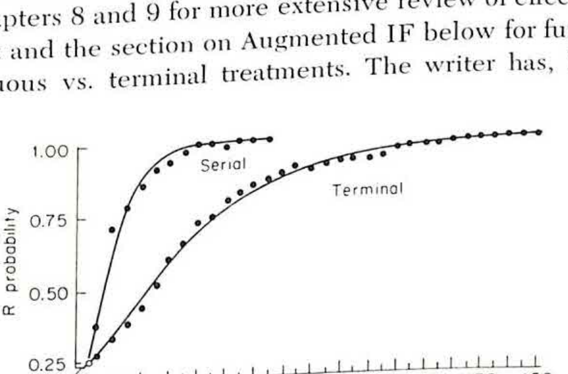


FIG. 2. Terminal and serial IF in trial-and-error learning. Probability of a correct response as a function of practice. (After C. E. Noble & Noble, 1958).

liberately omitted Chase from the above incomplete list of contributors to the literature on delay in repetitive and continuous responding. Though Chase and his colleagues have used some interesting devices, the work [fairly completely reviewed by Chase (1963)] generally suffers design, procedure, and analysis weaknesses far in excess of experimental psychology's tolerances.

If off-setting feedback by half a second would ruin rotary pursuit tracking (as it alternately scalds and freezes a showerer who fails to wait for one faucet turn to take effect on his water temperature before he makes another turn), one can say, "Why not?" as well as, "How much?" A subject who must respond to present error, given, say, an error signal jointly determined by present programmed input and past response output, must be able to separate the two, remember and anticipate program input, remember whatever manipulations he has made since the responses now being reported, and estimate the effects of his intervening responses. If the subject were perfect at these functions, tracking would hardly be a learning task for him. In a control-lag context Conklin (1957) concludes that any variable that reduces the accuracy of short-term prediction adversely affects tracking. And other evidence shows that a subject can neither keep track of past events (Pollack & Johnson, 1964) nor extrapolate target events perfectly (Battig, Voss, & Brogden, 1955b). The possible effects of pure delay are at least confounded with the incidental effects of interference and estimation; certainly any extension of fading-trace thinking is here misapplied.

## Locus vs. Delay of IF

Lorge and Thorndike used ball throwing and discrete trials in studying delay of IF, selecting task and procedures in order to "... discover the effect of sheer empty delay ... which interferes as little as possible with the belonging of the after-effect to the connection" (1935, p. 186). Their point seems well taken (though studies of delay in all kinds of situations are needed), and here a neutrally filled delay is accepted as the best practical substitute for absolutely empty delay—as the least complicated condition for testing the effect of delay *as* delay. Figure 1 shows the temporal events of such a single positioning trial; as Brown (1949) and later E. A. Bilodeau and Bilodeau (1958b) have pointed out, delay of IF is not the only time interval that might influence behavior. Indeed, there are three potentially effective experimental variables: IF delay—lag from R to its IF; post-IF interval—time elapsing between IF and the start of the next R; and interresponse interval (II)—time between successive Rs. It is obvious that if II is kept constant as IF delay is varied, then the post-IF interval varies inversely with IF delay, and any behavioral effect is more properly attributed to *locus* than to *delay* (E. A. Bilodeau, 1965), as the effect may as much depend upon the proximity of $IF_n$ to $R_{n+1}$ as upon its proximity to $R_n$.

More precisely, if any of the three intervals is constant, both of the others vary as one is varied, and an investigator who wishes to study them must seek ways of unconfounding their effects. One cannot, of course, without using reminder treatment, feed independent values of the three temporal variables into a single treatment, as any two determine the value of the third. But one can separate the effects of these confounded variables with any design in which the same two are not covaried in all groups, i.e., an experiment can have three sets of treatments, a different one of the three temporal variables constant from set to set (Bourne & Bunderson, 1963). If there is an expected optimal value of one of the three, that variable can be held constant at its presumed optimum, as the other two are covaried to determine if *either* of the others is indeed an effective variable (Denny, Allard, Hall, & Rokeach, 1960). And a reminder of the IF can hold post-IF interval constant at a presumably favorable value as delay alone or II alone varies (E. A. Bilodeau & Bilodeau, 1958b). For example, a reminder design using IF delay of 1, 5, and 30 seconds might repeat IF for all groups 15 seconds before the next R is due, to hold post-IF duration at 15 seconds as delay of IF varies in an interresponse interval of 45 seconds.

None of these illustrative studies, which, incidentally, were all designed to investigate the possibility that more than delay of IF might be operating, produced any support for delay of IF as an effective variable. The two studies using a positioning task supported II as the only temporal variable that mattered, locus of IF irrelevant (E. A. Bilodeau & Bilodeau; Denny *et al.*); Bourne and Bunderson found longer post-IF intervals beneficial in their concept-formation task. These differences do not seem to reflect an interaction of task and temporal variables, but rather a difference in the post-IF values studied. The two positioning studies used a minimum post-IF duration (8 – 10 seconds) about equal to the maximum used in the concept-formation task. Weinberg, Guy, and Tupper (1964), using a wider range of post-IF values, have recently reconciled the several post-IF outcomes; they found post-IF durations longer than 5 seconds equally good in a positioning task, and a 1-second post-IF treatment inferior to the others.

Some (not necessarily correct) reasoning to justify keeping track of the three intervals follows, to be followed in turn by a summary of some of the data that speak for taking all three into account. Parts A, B, and C of Fig. 3 present a most naïve and untenable case for predicting significant IF-delay effects on the basis of a straw-man fading trace theory. Part A presents the hypothetical decay of the trace of a response, $R_n$, that may be followed closely (arrow at $t_S$) or not so closely (arrow at $t_L$) by IF; if $R_n$ was correct, a connection is established at $t_S$, with a trace very like the trace of the correct R at $t_0$, and a less favorable connection is stamped in at $t_L$. (Similar, though less simple descriptions of stamping out can be offered when $R_n$ is reported as incorrect.) Presumably, then, the dominant tendency at the respective IF points would result in a more accurate $R_{n+1}$ at $t_S$, as in Part B of Fig. 3. But this simple-minded trace story is incomplete—the trace of $R_n$ is not the only event that can decay or be interfered with; IF, alternative Rs, and the new dominant tendency [as discussed and demonstrated by E. A. Bilodeau, Sulzer, & Levy (1962) and E. A. Bilodeau & Levy (1964)] are all at least potentially susceptible to forgetting. If one considers only the new R tendency, there seems no reason to assert that it is stable over the period from its establishment to its use. Part C, however, of Fig. 3 makes just this foolish assumption, with the necessary implications that the shorter delay after $R_n$ will yield better $R_{n+1}$ performance, regardless of time from $IF_n$ to $R_{n+1}$ and that $R_{n+1}$ will be insensitive to any temporal variable but delay of IF. A different hypothetical case is presented in Part D of Fig. 3, where error in $R_{n+1}$ increases with time between $IF_n$ and $R_{n+1}$—and where, depending on when the test is made (II constant), it is possible to find any outcome: delay irrelevant, short delay beneficial, or even long delay superior.

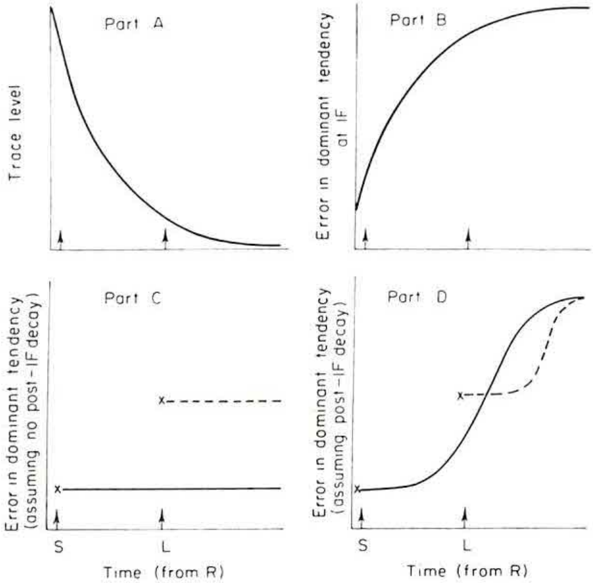


FIG. 3. Graphic representation of simple-minded, hypothetical trace decay with time after response (A); corresponding error in the dominant R tendency at IF (B); error in the dominant tendency after IF, assuming R tendencies change only before IF (C); and an alternative (D).

Hypotheses being free, we might speculate indefinitely about decay in the post-IF interval and its sources; extent; shape and slope, and the variables that affect them. The present purpose is merely to make a case for considering time *after* as well as time before IF as an interval in which events significant for the next R can occur. Forgetting possibilities are not the only ones pointing to research into both post-IF and IF delays. Bourne and Bunderson (1963) make a different case for the importance of the post-IF interval (in concept formation), showing that a short post-IF period has adverse effects on $R_{n+1}$. Interresponse interval makes its own case as a universal variable that one must take into account in any study involving at least two trials. One thing, at least, does seem clear: that Parts A, B, and C of Fig. 3 do not predict the outcome of actual studies in which time between $R_n$ and $R_{n+1}$ is varied. MacPherson, Dees, and Grindley (1948), and later E. A. Bilodeau and Bilodeau (1958b) and Denny, Allard, Hall, and Rokeach (1960), have all presented evidence that as the interval between successive positioning responses increases, accuracy declines. Bilodeau and Bilodeau further found the decline in accuracy to occur even when they combined a short delay of $IF_n$ with a reminder of the $IF_n$ shortly before $R_{n+1}$ to make both delay and post-IF intervals near optimal.

The fact is, of course, that there has never been much empirical support for a gradient of delay of IF in any task in which the subject learns a single R and in which the IF is given before the subject is obliged to use it. If the subject may delay $R_{n+1}$ in a series until he has received $IF_n$, it does not seem to matter if $IF_n$ is interpolated immediately after $R_n$ or later in the interval between $R_n$ and $R_{n+1}$. Lorge and Thorndike (1935) made similar statements when they reported insignificant delay effects in a ball-tossing experiment. Even studies of IF in classroom situations yield little support for immediate vs. delayed feedback when they are not taken at title value (Morgan & Morgan, 1935). Otherwise, through the mid-1950's there was just one prominent demonstration of a gradient of delay (more properly of *locus)* in a positioning task – Greenspoon and Foreman's (1956) study of line drawing, with II constant. The experiment was a replication, with better control of response origin and a change in the activity interpolated between Rs ("hand-maintaining" or not), of an earlier null study by Saltzman, Kanfer, and Greenspoon (1955). Portions of the Greenspoon and Foreman experiment have since been repeated, in line drawing, with a variety of modifications and controls by McGuigan (1959), E. A. Bilodeau and Ryan (1960), Denny *et al.* (1960), and Becker, Mussina, and Persons (1963). With delay and II covarying, McGuigan reported no delay effect in his major experiment. E. A. Bilodeau and Ryan used both "hand-not-maintained" and "hand-maintained" treatments and two of Greenspoon and Foreman's original *locus* treatments – 0-second delay–30-second post-IF and 20-second delay–10-second post-IF.[1] The main effects and interaction all failed to reach the .05 level of confidence. Denny *et al.* had their groups draw 60-glub (3-inch) lines, gave IF in glubs, and found no difference between groups with equal II and different locus of IF (0-second delay–30-second post-IF vs. 20-second delay–10-second post-IF) but a difference at the .02 level between two groups with immediate IF and different II (0-second delay–10-second post-IF vs. 0-second delay–30-second post-IF). Becker *et al.* varied delay, post-IF, and II, as well as omitting IF for one group, and found both in acquisition and in a retention test (without IF) that the only significant effect came from the presence of IF during acquisition.

In five studies with various positioning devices, and intervals ranging from 20 seconds to 7 days, E. A. Bilodeau and Bilodeau (1958b) also failed to find any trend in error with increasing delay, independent of the trend introduced by lengthening II (IF delay constant and short). Characteristic learning curves for short- and long-delay treatments can be seen in Fig. 4, where neither rate of approach to asymptote nor asymptote seems to depend on delay. Rounding out the above studies, which have minimal delays of 0–3 seconds and other delays of at least greater than 10 seconds, is a study by C. E. Noble and Alcock (1958) in which treatment differences in delay were

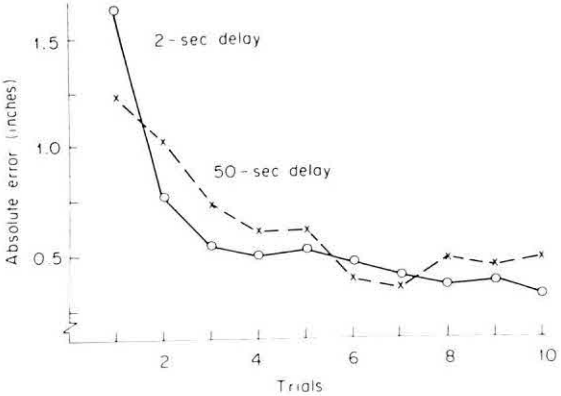


FIG. 4. Positioning accuracy and delay of IF. Error on successive 60-second trials for brief and long delay. (From data of Larré, 1961, with 0 interpolated Rs; $n = 30$ for each group.)

---

[1] The experiment repeated the treatment, 20-second delay–10 second post-IF, rather than the 30-second–0-second extreme of the original (a) because of the technical difficulty of *approaching* a 0 post-IF value and (b) since a post-IF period that approaches 0 gives the subject no time to process his IF. The 30–0 second combination, thus, uses unfavorable values of two variables, not just of delay.

measured in fractions of a second and the total range of delay was 0–3 seconds —again with discouraging results for proponents of the efficacy of delay.

In brief, there are several recent strong experiments reporting no delay effects in simple positioning (and other) tasks where IF is interpolated between the response to which it applies and the next response in the series. There are a few studies claiming (at the .05 level) delay effects without adequate experimental design. When delay and post-IF intervals are confounded and the latter is very brief, the findings of Bourne and Bunderson (1963) and Weinberg, Guy, and Tupper (1964) both support the inadequate post-IF interval rather than the long delay of IF as the source of such locus effects as have been produced (E. A. Bilodeau, 1965).

Ryan and Bilodeau (1962) obtained a marginal effect during countertraining in a lever-positioning task with brief vs. 7-day delay. Subjects first received a block of trials with brief IF delays and a constant target, and then responded once a week, the target shifting by a constant amount for successive Rs. The data are presented in Fig. 5, where it is clear that both groups shifted R magnitude, but not enough, and that the difference between groups is both small and constant from trial to trial. The experimenters were inclined to attribute their .05-level difference to chance; another possibility is that the delayed group took one trial to adjust to the change in IF treatment and suffered that much detriment throughout; and another is that the differ-

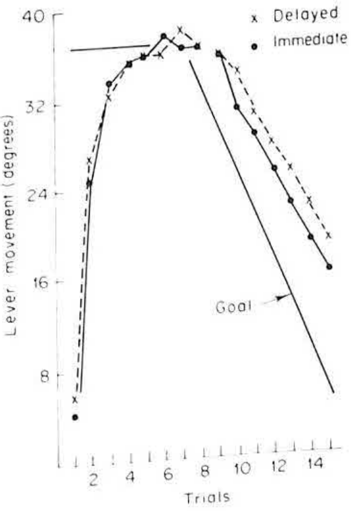


FIG. 5. Countertraining and delay of IF. Mean R amplitude on training trials (1–8) with immediate IF and on countertraining trials (9–15) with immediate IF or 1-week delay. Straight lines without symbols show the required R amplitude. (After Ryan & Bilodeau, 1962).

ence reflects the greater resistance to change in the original R that some writers have claimed for delay treatments, unfortunately for rather weak experiments (Lavery & Suddon, 1962; Brackbill & Kappy, 1962). Ryan and Bilodeau's experiment might be replicated by anyone who finds the results sufficiently suggestive.

Until recently the writer had been inclined to go along with the teaching machine propaganda that in verbal, problem-solving, and selective-learning situations delay might be important. The basis for this was an experiment by Saltzman (1951), who found delay of IF detrimental to verbal-maze acquisition and pointed to the differential opportunity to perseverate on errors as a possible determinant of the obtained differences. Additionally, of course, intratask interference in some verbal tasks (though not Saltzman's) with many Rs to be learned might produce incidental delay effects. But in recent years a sizable number of null results have been reported in all kinds of situations. Brackbill and her associates (1962, 1963, 1964) have a series of experiments in selective learning with children in which the typical finding seems to be a trivial, statistically nonsignificant, difference favoring short delay in acquisition, and a consistent suggestion (at about the .05 level) that delay improves retention. The retention differences may, thus, be direct or indirect effects of delay (better retention, the longer the training); but in any case, short delays do not produce dramatic benefits in selective learning (C. E. Noble & Alcock, 1958). Other nonmotor studies, too, show no delay effect—as, for example, in paired-associate learning (Brackett & Battig, 1963; and Kintsch & McCoy, 1964, with separate treatments controlling post-IF and intertrial intervals).

### INTERPOLATED ACTIVITY

Finally, of course, there are demonstrations that delaying IF can reduce positioning accuracy when the delay interval exceeds the interresponse interval. Lorge and Thorndike (1935) concluded that their subjects could improve with 6 seconds empty delay, but not when IF referred to the second previous, rather than to the immediately previous, throw. This trials-delay treatment displaces the successive IFs by one or more responses. Without trials delay the sequence of events is $R_1$-IF$_1$, $R_2$-IF$_2$, ... , $R_n$-IF$_n$; with a one-trial delay, the sequence is $R_1$, $R_2$-IF$_1$, $R_3$-IF$_2$, ..., $R_n$-IF$_{n-1}$; and with a three-trial delay the IF is displaced by three Rs; $R_1$, $R_2$, $R_3$, $R_4$-IF$_1$, . . ., $R_n$-IF$_{n-3}$. In two later experiments (I. McD. Bilodeau, 1956), it was shown that learning does take place despite trials delay, at least up to the maximum (five-trials) delay used, though error increases with delay. Lavery and Suddon (1962) have produced similar acquisition differences and have also tested for residual effects after IF is withdrawn; the retention effect, though not strong, favors trials-delay over immediate IF treatment. Bilodeau suggested that making other Rs between $R_n$ and $IF_n$, and the associated opportunity for interference, accounted for the trials-delay effect, and emphasized the analogous demands upon the subject in tracking with lag and positioning under discrete trials presentation, but with trials delay. Larré (1961), in turn, has shown that merely making positioning Rs is not the critical variable; he interpolated responses *without* IFs between responses for which IF was given, varying both the number of non-IF responses (0, 2, or 6) and their locus with respect to IF (after and with 2-second delay, or before and with 50-second delay). Neither main effect was significant, nor was there any evidence of interaction. Larré related his null result to the fact that his subjects (who knew that they would have some Rs with IF, and some without, and who were instructed to try for the target on all trials) held R amplitude constant over the non-IF trials, and changed R amplitude only on the R immediately after IF in all treatments. Confusing the R to which IF belonged would hardly matter. The subjects might, with equal effect, have applied $IF_n$ to any response between $IF_{n-1}$ and $IF_n$. Further investigation seems warranted if the sources of the trials-delay effect are to be specified.

## Frequency of IF

Frequency can refer to either relative or absolute numbers of IFs over a series of trials. From their investigation come the statements commonly made of the empirical properties of IF, that (a) there is no improvement in performance unless IF is introduced; (b) performance improves with IF; and (c) performance decays if IF is withdrawn (Thorndike, 1931; Lorge & Thorndike, 1935; Eaton, 1935; Elwell & Grindley, 1938; Lincoln, 1954; Becker *et al.*, 1963). Figure 6 provides representative data demonstrating these properties on a positioning task. This may be a good place to raise and dismiss the question of whether IF is necessary for learning, and whether R-produced kinesthetic cues [sometimes called intrinsic (Annett & Kay, 1957)] are IF. Presumably IF can be presented for any sensory channel, and an investigator who claims that he has reduced IF to zero may be wrong. Many devices have IF built in, and though more can be added, what one studies thereby is augmented IF, not the effect of the presence of IF. By our present definition, on the other hand, not all stimulation is IF.

Here IF is labeled a function of the discrepancy between $R_n$ and G; and IF means giving the subject either stimuli that are contingent on this difference or separate stimuli coding G (an index of the correct R on some scale) and $R_n$ (E. A. Bilodeau, 1955a; Suppes, 1960; Rouanet & Rosenberg, 1964). Response-produced stimulation, important though it is in performance (Bahrick, 1957), is not IF unless it satisfies these requirements.

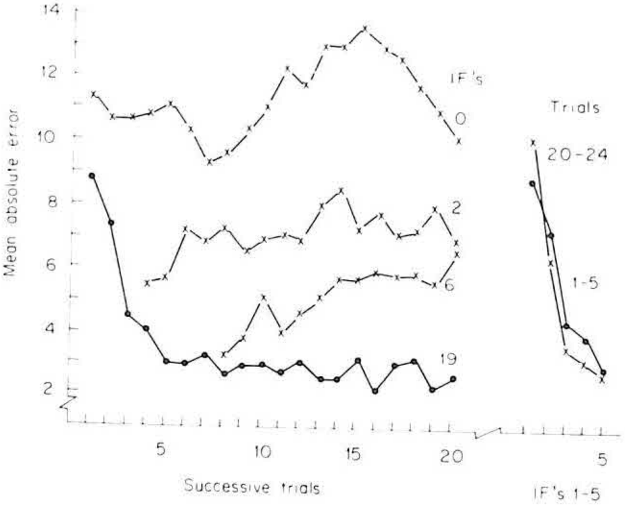


FIG. 6. The effects of introducing and withdrawing IF early and late in practice. Mean error on trials with and without IF. Group 19 had IF on all trials, Group 0 did not receive IF until after Trial 20, while IF was withdrawn after Trial 2 for Group 2 and Trial 6 for Group 6. The right-hand portion presents the first five trials with IF for Groups 0 and 19, to demonstrate the absence of latent learning. (After E. A. Bilodeau *et al.*, 1959).

The kinesthetic stimulation accompanying a manipulation is not IF unless some standard is also provided, or some function of the difference between actual and required Rs is expressed kinesthetically. Lincoln (1956) has tried both on a handwheel and compared them with verbal IF. The group given the *kinesthetic standard* (handwheel driven at the correct rate for 15 seconds after every 15-second trial) was poor and erratic, while the group given *kinesthetic error* (handwheel driven at a rate equal to the mean error in rate on the last trial) stayed very close to the group with verbal IF. As Lincoln emphasized, the kinesthetic-error treatment also used verbal IF for direction of error (which by itself can be as effective as more complete IF), so that the difference between the kinesthetic-error and kinesthetic-standard treatments has alternative interpretations.

To the writer it makes sense to maintain that learning in positioning means building a scale relationship (or a new scale) between kinesthetic stimuli and some other stimuli—meaning perhaps very much what Bahrick, Fitts, and Schneider (1955) intend when they refer to using knowledge of results to improve the subject's use of kinesthetic feedback. The subject may attach different numbers to different kinesthetic values, translate the kinesthetic values into visual extent, auditory intensity, etc. Unless, however, he (happily) transfers an old scale that fits the new situation, he needs IF in order to make the appropriate present transformation — even if he does know, blindfolded, that he has moved his arm 23.04 inches and through 11.16° of arc. (In fact, of course, many people have no idea within yards of how far they have moved arm or pencil, and when they do we are more inclined to refer to past overlearning than to present learning based on R-produced cues.)

Among the immediate flood of questions, problems, and inconsistencies that arise in the attempt to narrow down what we mean by IF, its presence, and its complete absence, three seem most important. (a) Another investigator might include visual, auditory, olfactory, etc., cues that accompany behavior as R-produced, and, by his definition, find learning possible on R-produced feedback. The writer agrees to call these R-produced stimuli, however, only when we are concerned with seeing, smelling, etc., and not when we are studying limb movements. Seeing, smelling, etc., do show learning improvements when IF, defined by other external stimuli, is introduced (Hamilton, 1929; Alluisi & Sidorsky, 1958; Larimer & White, 1964; Atkinson & Carterette, 1964; Foley, Lavery, & Abbey, 1964). Vigilance behavior, too, is modified by using other stimuli as IF (Mackworth, 1950; Chinn & Alluisi, 1964). (b) Although Lincoln's study of kinesthetic IF in handwheel turning (1956) leaves their effectiveness open to question, there are other claims that kinesthetic stimuli may work. Sulzer (1961), for example, put an artificial block on a lever device to stop the response at the goal, and found alternating these trials and trials without IF about as effective as an equal number of standard verbal-IF trials. Blick and Bilodeau (1963), on the other hand, gave various directions and amplitudes of guided Rs between positioning trials without finding any evidence of either positive or negative transfer of guided training. [And, does guiding the subject to the goal mean giving him IF, a trial of guided practice that transfers to the standard task, or both? The reader is directed to Holding's (1959) discussion of IF and kinesthetic cues in guided tracking. If guided practice removes, or at least limits, IF, the limiting may be more detrimental to complex than to simple tasks.] Elaborate tests of kinesthetic IF possibilities in positioning tasks, in brief, can be expected to raise more dispute than they will settle on a short-term basis. (c) Pretraining tuitions can benefit performance — providing one with useful information that is not, strictly, feedback (E. A. Bilodeau & Bilodeau, 1961). It would be too easy to fall into the paradox of insisting that the information aspect of IF is what counts, while maintaining that only information feedback can produce learning. No doubt the benefits of guided practice would be about equal in a free-R test, if one guided-training sequence were score$_n$ − R$_n$, and the other R$_n$ − score$_n$. This is admitting that, other things equal, the order in which information and R pairs occur may sometimes be trivial, but as the usual order is R-IF, information feedback is at least ordinarily necessary.

Finally, of course, it is improper to claim that only learning produces R modification, or even all reduction in error. Adaptation effects might modify the behavior in ways that would lead to consistent error reduction. There are examples of improved adjustment to the vertical (Solley, 1956; Pearson & Hauty, 1959, 1960), for which the writer can neither specify the IF nor offer a satisfactory account in adaptation terms. And in some highly overlearned tasks the subject may rely upon previous learning rather than present IF to control his Rs; warm-up under these conditions might be mistaken for learning.

The data plotted above in Fig. 6 present one kind of evidence of the effect of absolute frequency of IF on performance and learning. (a) With succeeding IFs, error is progressively reduced toward a minimum value (i.e., performance improves). (b) Residual (learning) benefits are greater when IF is withdrawn (paralleling extinction) after the larger number of pretest IFs. Any of other numerous tests without IF after some practice with IF can be used to make the point that learning does take place in the usual R-IF conditions of practice (Thorndike, 1931; Lorge & Thorndike, 1935; Lincoln, 1954; Annett, 1959; Karlin, 1960; Lavery & Suddon, 1962; Lavery, 1964). Additionally, Fig. 6 shows again (Eaton, 1935; Hamilton, 1929) that a first series of Rs without IF does not produce latent learning.

Other data bearing directly upon the influence of IF's absolute frequency (total number of Rs followed by IF) and relative frequency (ratio of IFs to responses) on learning and performance have been obtained by giving occasional IFs in a series of responses. In a positioning study in which subjects turned a knob, IF was given (fixed ratio) for $R_1$ and thereafter for every R (100 percent), every third R (33 percent), every fourth R, or every tenth R (E. A. Bilodeau & Bilodeau, 1958a). The total number of Rs was adjusted from group to group so that every subject received 10 IFs. The effect of absolute frequency is represented in the left half of Fig. 7, where only those Rs immediately after an IF are plotted, i.e., absolute number of IFs is constant, relative frequency is variable. With absolute frequency constant, there are no differences to be attributed to relative frequency, and the experimenters concluded that absolute frequency was the learning variable, and relative frequency irrelevant, at least up to a 10 percent ratio. Part of the data are replotted in the right half of Fig. 7 against ordinal number of R, rather than IF, to support this view. The 100 percent treatment yields a continuous curve of error reduction; the partial IF treatments yield discontinuous, step functions. Trials without IF make little change in R (or increase in error), change occurring between Rs for which IF is given and the next R in the sequence. A later study by Larré (1961), previously discussed under Locus (p. 267), allowed IF for every R, every third R, or every seventh R, and verified these

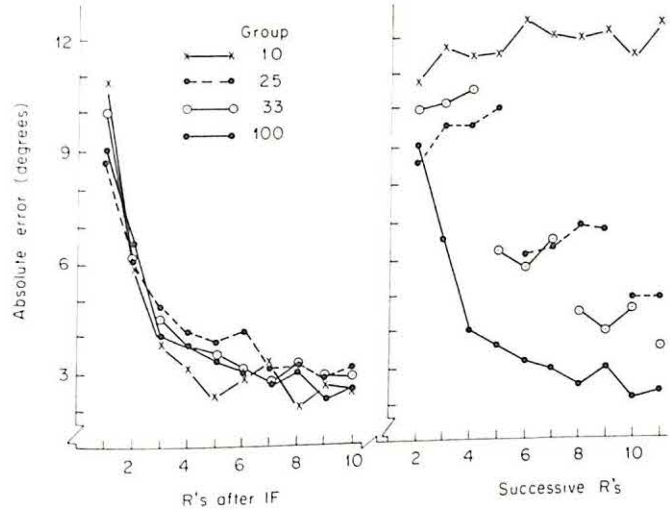


FIG. 7. The effects of relative and absolute frequency of IF. On the left, mean error in Rs immediately preceded by IF (absolute frequency constant, relative frequency variable). On the right, mean error in Rs 2–11, breaks in the curves indicating IF (relative and absolute frequency variable). (After E. A. Bilodeau & Bilodeau, 1958a).

findings. (If number of Rs is constant, performance depends upon relative frequency of IF; if absolute number of IFs is constant, relative frequency does not matter.)

Bourne and Pendleton (1958) in concept identification, and A. Taylor and Noble (1962) in selective learning, have varied relative frequency of IF and found that R error over a block or series of Rs was greater the fewer the IFs within a block. At least in the early stages of the learning tasks so far explored, there is no evidence that other cues have served as secondary IF, as secondary reinforcement can substitute for primary (Denny, 1946). On the contrary, Rs not followed by IF show either no trend, or a negative trend. The absence of secondary effects may imply a greater success in reducing outside, supplementary cues to a minimum for human than for animal subjects.

Roughly analogous techniques in more complex tasks, outside the scope of this chapter, also show losses in proficiency when IF is removed or given intermittently (Lincoln, 1954; Battig, Voss, & Brogden, 1955b; Poulton, 1957; Gordon, 1959). The reader is referred to Chapters 8 (pp. 374–375) and 9 (pp. 431–433; 472) for coverage of these important tasks and their important frequency-related variables.

Also outside the scope of the present chapter is the most extensive body of research of all treating frequency and frequency schedules in human learning: probability learning, in which frequency of IF is the direct (and almost sole) concern in a vast experimental literature, and an almost equal theoretical literature. Probability learning studies provide some of the clearest illustrations that IF, usually called reinforcement in this context, is arbitrary, and that the dominant R and its degree of dominance depend upon absolute frequency, relative frequency, and pattern of reinforcement. This crude summary, though it suits our present purpose, does gross injustice to the scope of probability learning and the precision for which the several mathematical models strive. Fortunately the reader can refer to Estes' recent review and Anderson's comments in *Categories of Human Learning* (Melton, 1964), and to Noble, Chapter 2 (pp. 52–55) of the present volume, for a comparison of probability and selective learning.

## Augmented IF

Supplementary, augmented, or extra IF as used here merely means any IF added to a standard task, with no intended implication of a distinction between IF intrinsic to the task in the sense of being an unavoidable consequence of responding (say kinesthesis), and IF extrinsic to the task in the sense of not being a direct and necessary consequence of responding [a fairly common contemporary view; see Annett and Kay (1957) and Fitts (1964), for example]. All our IFs are extrinsic in this sense, whether extra or not. For example, if IF usually follows R, and a new treatment introduces IF while R is in progress, the new treatment has augmented IF; if a similar device ordinarily presents visual IF throughout the course of responding, an extra-IF treatment might add auditory IF to the visual. From this point of view there is no fundamental or natural IF standard, and, further, another name might serve the present topic as well as the one used here. With any manipulation of time, place, or amount, augmented IF could be treated as an IF transformation, with reduced IF on the other side of the standard. Or, as the major interest in extra IF is what is learned, and as tests for residual effects are emphasized, the present section might be relabeled transfer of training.

While it is present, extra IF has been found to improve positioning accuracy, and either to improve (Williams & Briggs, 1962) or to have little effect on tracking performance (Archer, Kent, & Mote, 1956). Transfer of training is positive, evaluated by the A-B vs. – -B design. The usual question, though, is not whether augmented-IF habits transfer to a standard task, but how this transfer compares with the benefits of training on the standard task itself. The A-B vs. B-B design is common, with the groups compared on performance in Test B; or, in positioning, in the absence of any IF. All three possible transfer outcomes are on record: (a) test performance worse after training with extra IF than after standard IF (Miller, 1953; Karlin, 1960); (b) test performance about the same (Stockbridge & Chambers, 1958); and (c) test performance better after augmented IF (Smode, 1958). Logically, of course, extra IF might have any of these effects, i.e., might (in habit terms alone) (a) serve to sustain performance while an interfering habit develops; (b) have an additive performance effect, establishing an independent habit; or (c) have an incremental effect on the standard habit, by allowing more adequate interpretation of the standard IF. Some of the most diverting divergences in experimenters' opinions can be enjoyed in the literature on extra IF. For example, E. A. Bilodeau (1952d) found little residual effect of extra IF in a standard test, and explained his findings in habit and cue terms; Smode's (1958) differences did carry over to the transfer test, and he favors a motivational interpretation.

### Positioning Rs

Studies of positioning Rs show performance benefits in training from extra IF (Annett, 1959), and less benefit to the standard task than from an equal amount of standard-task training. The uniformity of the outcomes for positioning tasks no doubt reflects the communality of the treatments across studies: a standard task providing IF at the end of the response and an extra-IF task giving IF while R is in progress. Much of the work is highly compatible in intent with research on shifts from exteroceptive to proprioceptive R control as a function of level of learning in more complex tasks (Bahrick, Noble, & Fitts, 1954; Bahrick & Shelly, 1958). Miller (1953) uses *learning* and *action* feedback for the respective terminal and time-shared treatments. The present writer likes *action* IF, as descriptive of the treatment, but prefers to substitute *terminal* for *learning* IF. In serial, selective learning C. E. Noble and Noble (1958) make a comparable distinction between serial and terminal IF, depending upon whether the subject gets IF after every choice or only at the end of a series of choices; as Fig. 2 shows, serial makes for faster learning. Terminal IF represents a kind of delayed or displaced IF, of course, and may also be *lump-sum* (a single index for many components or items), though terminal IF may be given separately for every item.

Transfer tests after action IF is withdrawn show that subjects do learn with action IF (Lincoln, 1954; Annett, 1959; Levy, 1961; Fox, 1962). In tests without any IF, performance after action IF is inferior to performance after terminal IF (Lincoln, 1954); and tests under terminal IF show terminal-trained groups superior to action-trained (Fox, 1962). But transfer is positive from action to terminal IF conditions, and is an increasing function of the number of preshift action-IF trials (Levy, 1961). Accuracy of responding is constant over action-IF trials, of course, but this does not mean that a different measure would have no trend; a temporal measure might, also, show that transfer from terminal- to action-IF treatments is less than complete.

The contrast of the terms *action* and *learning* IF, though they are not the writer's preferred pair, can be readily defended as making the point that much of motor learning does not consist in acquiring precise R tendencies, but in correcting for present discrepancies between R and G, or, in positioning, that the subject will not acquire a precise R unless forced to. This is the same point made in tracking studies, where the skill does not involve reproducing an exact pattern of Rs, but rapidly correcting error in ongoing behavior—as shown by the very high transfer of training between different target patterns (Battig, Nagel, Voss, & Brogden, 1957). The roles of present error signal and past practice have been studied in positioning by changing the goal from one trial or practice block to the next (E. A. Bilodeau, 1953a; Suppes & Frankmann, 1961), with present error signal of paramount importance. Sulzer (1961) has taken a different approach to what is learned in positioning, using the lever described in detail in Chapter 7 (p. 320). Sulzer had two kinds of terminal IF (verbal score and a mechanical stop to guide the subject to the correct position) and tested for place vs. distance learning, asking if the subject learned to stop at a certain point or to move his arm through a given arc. In training, arc and place were constant, but in the test the starting place varied slightly. Under Sulzer's conditions, place learning was supported, the arc varying with change in starting position; this work, of course, needs following up with modifications to test whether action IFs change what is learned.

We probably also need some demonstrations that not every supplement to the standard IF benefits positioning accuracy, though we may wait for these, as experimenters are usually reluctant to set about producing null results except when very hard pressed. The present course, to vary pattern of alternation of terminal, action, and non-IF trials with a view toward determining what is learned and whether pattern of alternation can affect phasing of proprioceptive control, will more likely be pursued.

### TRACKING

In tracking tasks there are instances of both null and positive effects for extra IF while it is present. Where there have been training benefits they have sometimes shrunk when the extra IF has been withdrawn, and have sometimes remained in a test under standard IF. (The reader is referred to Chapter 8, pp. 400, 401, for further review of augmented IF and tracking.) There is, thus, some disagreement in interpretation; that is, there are two points to consider: (a) whether extra on-off target indicators are universally effective while they are present, and (b) the residual benefits of a previous extra-IF treatment, *before* the arguments about extra IF as cue, reinforcement, or motivation begin. [For reviews of the various findings and interpretations the reader is directed to Smode (1958), E. A. Bilodeau and Bilodeau (1961), and Adams (1964).] Extra IF can probably be redundant beyond the limits of providing any performance benefit, and be accepted as such regardless of whether the given investigator takes IF's role as cuing, reinforcing, or energizing; no one expects limitless benefits from adding to any variable. And some studies have failed to find any benefit from the presence of an extra IF (Archer & Namikas, 1958; I. McD. Bilodeau & Rosenquist, 1964) or even a slightly adverse effect (Archer, Kent, & Mote, 1956). Annett and Kay (1957) offer sight vs. touch typing as an example of short-term advantage for augmented IF against long-term advantage for standard training—an illustration of the advantages Bahrick *et al.* (1954) expect when simultaneous tasks depend upon separate sensory channels.

Unless IF is independently scaled, however, claiming that augmented IF benefits performance to the extent that built-in IF is poor is circular, even if correct. Adams (1964) also makes a case that the learning benefits of extra IF are dependent upon the usable IF in the standard task. We need studies varying quality and quantity of the standard IF—along the lines Poulton (1957) used to impoverish IF, by using noise to impoverish built-in IF (Kinkade, 1963), or following systems engineering in designing good and poor displays—and variety, quality, and quantity of extra IF. If subjects performed on a Standard Two-hand Coordination Test, but gazed at the apparatus through a progressively more opaque glass screen from group to group, extra auditory, temperature, visual, or other IF should be increasingly beneficial as the screen darkened. The writer would wager that augmented IF would have no benefit, however, without the screen's degrading effect on built-in IF. The benefits of a passenger's serving as an extra source of information likewise depend upon the conditions of the present drive and the demands upon the driver's attention. Under optimal conditions the extra IF has only annoyance value, but in dark, storm, and heavy, fast traffic it can be a welcome aid to performance, for example, by supplementing standard IF from exit signs.

The Pedestal Sight Manipulation Test (PSMT), a gunnery trainer adapted from USAF World War II equipment, is one device that consistently yields large performance effects to augmented IF. Three R components are usually required. The subject swings the sight head vertically and horizontally to keep a pipper on the nose of an image of an aircraft moving across a screen, and turns a knob to adjust a projection of the range ring to the size of the target's wings. The factors (division of attention, ambiguity, and contradiction) that make the PSMT particularly susceptible to augmented IF have been discussed elsewhere (E. A. Bilodeau & Morin, 1951). R. H. Seashore, Underwood, Berks, and Houston in an unpublished report (Underwood, 1949) began the PSMT work by adding a filter to turn the target image red when the subject was on target. As long as it was present, this extra on-target indicator helped performance, but later studies with various augmenting treatments found much of the benefit lost when the filter was removed either in rapid alternation of extra- and standard-IF trials or after a block of trials (Houston, 1947; Morin & Gagné, 1951; E. A. Bilodeau, 1952d; Goldstein & Rittenhouse, 1954). Further, in an experiment using procedural controls to prevent the subject's learning the idiosyncracies of a particular unit, differences between standard- and extra-IF treatments were shown to grow at the same rate whether the extra IF was added early or late, to suggest that the filter develops a second and independent habit (E. A. Bilodeau, 1952d). Some of this research was done with an eye on augmented IF for training programs, and from the applied position the outcome is far less interesting than from the pure. Some of the difference in emphasis given the small and diminishing residual extra benefit of filter treatments also no doubt relates to the applied vs. pure emphasis of the different studies.

Other devices have produced similar training effects—probably more positive than negative instances—(for example, Hauty & Payne, 1955; Smode, 1958; Williams & Briggs, 1962; Kinkade, 1963), with residuals present in the postshift test under standard IF; a start has been made on studying pattern of extra IF, comparing on- and off-target indicators, and identifying variables controlling the benefits of augmented IF and transfer to the standard task. Combining augmented IF and component analysis also brings out a few points. Giving IF for one component of a multidimensional task may not improve criterion performance by one weighting of components in a composite score, but it can affect the profile of component proficiency (E. A. Bilodeau, 1955b); if relative proficiencies change, then a change in weighting could make a component supplement effective indeed. Chinn and Alluisi (1964), *introducing* IF into a vigilance task, have a like demonstration, that both what the IF covers (missed signals, correct Rs, false Rs) and the R measure (missed signals, false Rs, latency) must be considered in describing IF's effectiveness.

Clearly, determining both (a) the conditions under which augmenting IF affects performance and (b) the variables controlling residual (learning) benefits from extra to standard IF are important. Very few investigators have looked for more than a difference between a pair of treatments; there are not very many experiments of any kind with augmented IF; and there is a real need for obtaining relationships between continuous experimental variables and performance and learning effects of augmented IF.

## IF Transformations

It has already been redundantly stressed that all IF studies could be treated as various kinds of transformations, but what remains seems sufficient. These are largely *numerical* and *spatial* displacements, variation in the number reported to the subject as his score for a trial and in the location of light, pointer, or target follower on a spatial display. From the manipulation of the contingency of IF on R comes the most direct evidence of IF's important empirical property: that R and R change depend upon the magnitude and direction of error signaled in IF. Rate of acquisition or limit of accuracy depends upon the error transformation, and, as countertraining techniques show (see Fig. 5), IF can be used just as easily to steer the subject away from an established R as to establish the R in the first place. Transformations permit $IF_1$ to varying independently of the value of $R_1$, an advantage over the usual confounded measure of the joint effect of IF and R.

Two points are worth noting. (a) The relationship of $IF_n$ to $R_n$ (i.e., to the experimenter's measuring and scoring R) is arbitrary, and need be neither 1 : 1, nor even simple. To claim that IF is a function of the difference between the R executed and the R required is to admit that IF can be *any* function of R and that the *f* itself can be varied and its effects on behavior studied—the main interest of this section. An obvious illustration is the experimenter's freedom in choosing sensory modality and absolute stimulus values. For pulling a lever, turning a wheel, or pushing a switch, IF may come by dial and pointer, spoken words, or tones through earphones. What dial and pointer are used, the code for the words spoken, which ear will receive the undershoot signal, etc., are also arbitrary. In tracking, the same sequence of control movements can be required for either compensatory or simple following pursuit. (b) $IF_n = f(R_n)$ does not mean that the IF is the score recorded. In the various studies below, the scale on which R is measured (recorded score, R measure) is constant for all treatments. The experimenter rounds in measuring R for his records (for example, rounds a reading of 3.127 inches to 3.13, 3.1, or 3 inches), but rounds all groups to the same tolerance, measuring R before, not after, transforming. The IF may be "a hit," but the R recorded is the same −13° that would be taken for the same R in a treatment where IF is −13°. Otherwise *scores* may well differ from group to group, either because the groups act alike but get different scores for the same behavior, or because the actual behaviors differ (E. A. Bilodeau, 1952a, b, 1954, 1955a, 1956; I. McD. Bilodeau, 1965).

In positioning it is easy to make this scoring-performance distinction. In tracking where IF transformations are one portion of the investigation of control and display variables and their interaction, the distinction is also routinely made. In adapting apparatus not built with IF transformations in mind, however separating the IF grain from the grain of the R measure is not always a matter of course; a change in the size of the rotary pursuit target, say, often carries with it electronic implications that the same R can accrue time on target in one treatment but not in another. Scoring and R measures need not be confounded, but more than simply changing the size of the target disk is required (R. B. Ammons, Ammons, & Morgan, 1958).

Nonetheless, a reference to Chapter 8 by Poulton and Chapter 9 by Smith will show that the quantity and variety of IF transformations in continuous responding exceed the output for positioning, and will support the present contention that tracking transformations can be well thought-out and well executed. K. U. Smith and his associates, for example, have studied displaced IF of many varieties in continuous responding, measuring R independently of its representation to the subject (W. M. Smith, Smith, Stanley, & Harley, 1956; W. M. Smith, McCrary, & Smith, 1960; K. U. Smith, Wargo, Jones & Smith, 1963) – i.e., the R measure is taken from the track of the subject's pencil through the maze, not from a photograph of the TV screen that provides IF. And most others have varied and measured control and display features without confounding (Andreas, Finck, Green, Smith, & Spragg, 1959; Battig, Nagel, & Brogden, 1955a; Eason, 1963; Fitts & Deininger, 1954; Gordon, 1959; Hartman & Fitts, 1955; J. R. Simon & Simon, 1959; F. V. Taylor & Birmingham, 1948).

Incidentally, in measuring positioning Rs, it is probably wise to record R amplitude (or error) more finely than to two or three categories. The grounds for this preference are much the same as those making time-on-target low in the hierarchy of R measures in contemporary tracking research. Though reporting only "hit" or "wrong," one might score degree of error; the difference between a hit and a failure may be far less than between a pair of hits or two failures, and the kind of error may be as significant as frequency of hits (Dees & Grindley, 1951).

## POSITIONING

Manipulations of the relation of IF to amplitude of positioning Rs are divided into three categories: (a) scale-grain, or discrete, transformations; (b) continuous IF-R transformations; and (c) goal variability. Though most of the variations in the IF-R relation have been numerical, there have been some of the spatial displacements typical of tracking displays. The distinction bebetween spatial and numerical transformations has not been critical under the terminal IF treatments so far studied.

### *Scale Grain*

The scale-grain or rounding-error treatments are most easily summarized.

As IF is rounded from a score including a fairly precise statment of amount and direction of error, error increases, but the behavior is not very sensitive to coarseness of the scale grain. In any situation in which a continuous scale is used, there is some rounding, though the experimenter has the option of rounding to the nearest decimal, whole number, multiple of 10, etc. (E. A. Bilodeau, 1954, 1956). The experimenter must decide the tolerance (target size) he will allow in signaling a hit, and also the number (coarseness) of the other possible IF values. In any case, the relation of IF to R value is stepwise and discrete, though the size of the steps may vary, and may not be equal over the entire range. Figure 8 presents several scale-grain transformations of IF to R scale in which the tolerance about target center is kept constant; though they were used in a tracking study, they might also be used in positioning. The steps may be unequal (A); intermediate, and with a smaller step about target center (B); uniform and small (C); continuous, or nearly so (D); or any size and arrangement that may be desired. The R-scale, or true-error measure, of course, is unaffected by the error displayed to the subject.

Trowbridge and Cason (1932) early demonstrated that amount and direction of error had more effect than "right" and "wrong" only. Rounding IF (error reported to the subject in turning a micrometer knob, G = eight full turns, or 200 units) to the nearest unit, multiple of 5, 10, . . . 50 units, E. A. Bilodeau and Rosenbach (1953) found adverse effects for rounding errors only early in practice, and only for step intervals of 40 and 50 units. Data from the same task (turning a micrometer knob without visual cues), from an unpublished study by the writer can serve to illustrate both the effect and lack of effect of extreme treatments. One group received magnitude and direction of error (MD), another magnitude (M) only, and the third direction (D) only. All groups improved, with D as good, in this case, as MD, while M lagged. One can conclude that it is not always necessary to refine IF to the limit to produce optimal positioning accuracy, and, although there seems to be no example of this at present, it may be that IF could be refined to the point of detriment.

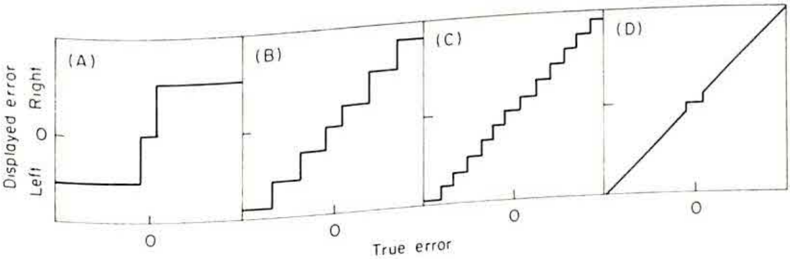


FIG. 8. IF transformations. Displayed error as a function of R error, with target size constant. (After Hunt, 1961).

*Continuous Transformations*

Target size or the range of Rs signaled correct disappears as a separate topic in the present treatment. Target size falls partly under scale grain, i.e., transformations in which either the size of the plateau at G *and* the number of other steps vary, or only Rs reported as hits are transformed [functions such as (D) in Fig. 8]. It also comes partly under reduced-error (RE) treatment (E. A. Bilodeau, 1955a), an example of which is shown in the insert of Fig. 9, with its related opposite, an IE (increased-error) treatment. The latter has one discontinuity in the IF scale, but is included as a continuous transformation anyway. In an RE treatment IF is expressed as

$$\mathrm{IF}_n = (R_n - G) - k, \tag{3}$$

when $(R_n - G) \geqslant k$,

$$\mathrm{IF}_n = (R_n - G) + k, \tag{4}$$

when $(R_n - G) \leqslant -k$, and

$$\mathrm{IF} = 0 \tag{5}$$

when $|(R_n - G)| \leqslant |k|$,

and in an IE treatment as

$$\mathrm{IF}_n = (R_n - G) + k, \tag{6}$$

when $R_n \geqslant G$, and

$$\mathrm{IF}_n = (R_n - G) - k, \tag{7}$$

when $R_n \leqslant G$.

In RE, thus, several different Rs will yield the signal of 0 error, and any error signal is reduced by up to $k$ units, while in IE, no R can be signaled a hit, and IF exaggerates any R error.

Perhaps the best demonstration of the consequences of inflating and reducing error is the data plotted in the left portion of Fig. 9, showing error in turning an unseen micrometer knob ($G = 200$) for size and direction of error transformation. With transformations up to $k = 20$ (20 percent of G amplitude, in target width), neither treatment affects absolute error, but both functions are positively accelerated and accuracy declines with further distortion, particularly for RE treatments. The variance data, in the right half of Fig. 9, are more interesting than the means. The estimated $\sigma^2$ (of average R on the last three trials) between subjects (est. $\sigma^2_{BS}$) is a sensitive measure of RE, but

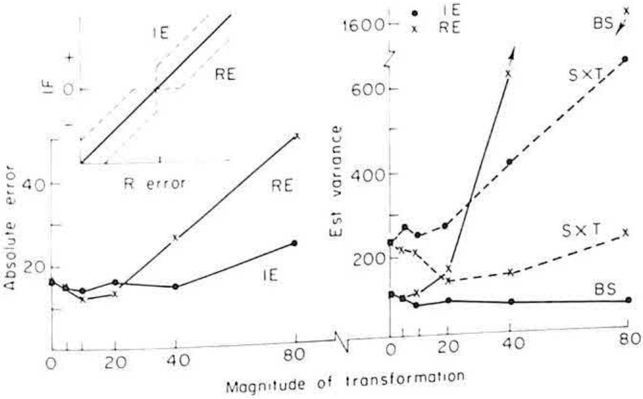


FIG. 9. Increased- and decreased-error transformations. Mean error and estimated variances as functions of magnitude of the increased-error (IE) and reduced-error (RE) transformations shown in the insert. (After E. A. Bilodeau, 1955a).

immune to IE, while the reverse is true of the estimated $\sigma^2$ of the subjects-by-trials interaction (est. $\sigma^2_{S \times T}$). Both make excellent sense. A larger target tolerance allows individual subjects to develop constant errors – the IF signals "correct" for almost anything, which may by chance be almost any R from subject to subject; and, as IF signals "correct," any subject has no reason to change his R from trial to trial (little $S \times T$ interaction). A big IE constant, on the other hand, discourages both constant errors and R repetition; IE always carries an $IF_n$ to which a change in behavior from $R_n$ to $R_{n+1}$ is appropriate, and, moreover, to which a change in the sign of the R error is appropriate.

To account for R insensitivity to target size over fairly extreme RE transformations in lever positioning, it has been proposed that generalized R tendencies (i.e., the tendency to repeat an R that has been called correct) about G can summate to produce essentially equivalent accuracy for different RE treatments (E. A. Bilodeau, 1954). Reasoning that the summation might not work to the same advantage if a different R were required on every trial, the writer tried and failed, in an unpublished experiment on the same lever, to show an interaction of two RE values and constant (C) vs. variable (V) goals over trials. The V subjects were told what goal to try for on each trial. The C subjects had the same goal on all trials, different subgroups representing the goals used for the V subjects, so that all goals were represented in both C and V treatments on every trial. Treatment V had about twice the mean absolute error (deviation from the goal assigned) of Group C at the end of the 13 trials, but size of RE (0 or 3°) had no effect on either V or C treatments.

Other transformations combine both continuity of IF = $f(R)$ and constant target size. Examples in turning a micrometer knob are linear transformations, IF = $mR + b$, in which IF (reported score) increases with R magnitude, with $G$ constant and slope ($m$) and IF intercept ($b$) covarying (E. A. Bilodeau, 1953b); the curvilinear transformation, IF = $aR^b$, that C. E. Noble and Broussard (1955) used to hold intercept constant as slope varied; and Schumsky's (1959) positive and complementary negative linear transformations (IF = $-\,mR + b$). The data in Fig. 10 well illustrate the effects of linear transformations on R accuracy and rate of acquisition; for either of two goals, rate of approach to the limit is affected by the transformation IF = $mX + b$, very little by the value of $G$. Though countertraining and incentive motivation interpretations can be offered for the same sets of data (C. E. Noble & Broussard, 1955), the nice dependence of the behavior upon the IF is clear. Trial 2 is most critical for relating $\mathrm{R}_{n+1}$ to $\mathrm{IF}_n$, as these studies had all subjects try to "make a practice R" of a given amplitude for $\mathrm{R}_1$, in order to reduce the influence of $\mathrm{R}_1$ upon $\mathrm{IF}_1$. The effect of magnitude of IF is shown, for example, in the $r = .70$ between $\mathrm{IF}_1$ and $\mathrm{R}_2$, as compared with $r = .10$ between $\mathrm{R}_1$ and $\mathrm{R}_2$ in a reworking of E. A. Bilodeau's 1953 data; the importance of direction of error signaled in IF appears in the decrease in amplitude from $\mathrm{R}_1$ to $\mathrm{R}_2$ that Schumsky produced by giving smaller scores for larger Rs. These micrometer transformations all slowed learning relative to the IF = 1.0R + 0

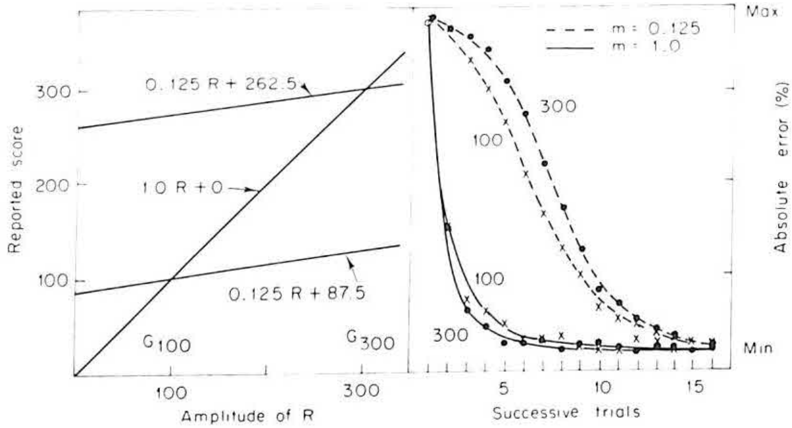


FIG. 10. Linear transformations and goal magnitude. Slope and intercept have been made independent by changing the required amplitude of the goal R. The right-hand portion plots error over trials, for goals and transformations. R-error scales have been arbitrarily adjusted to the same scale, percentage error over the obtained range for the respective goal magnitudes. (From an unpublished experiment by E. A. Bilodeau, 1952).

control, though asymptote was not affected; and, as Denny *et al.* (1960) made clear in using glubs rather than inches, slowing learning is often the point of a transformation. But not all transformations are expected to be detrimental, as witness M. E. Noble & Bahrick's (1956) choice of the transformation of pounds of pressure on a pressure stick to $IF = 4(R - G)$ units. Further manipulation of continuous transformations can be expected to yield facilitating as well as detrimental effects. Sulzer and Levy have such an example in an unpublished exploration of absolute size of IF in lever positioning, with $IF = m(R - G)$, $m = 1, 10, 100$, or $1000$ and the magnitude of $G$ not specified to the subject (1961). Only the $m = 10$ treatment differed even numerically from the others (was slightly more resistant to countertraining), but there are other demonstrations in levering that the sheer size of the number used for IF has some bearing on R change. An arbitrary, preassigned $IF_1$ value influences $R_2$, even when the subject is instructed to repeat his $R_1$ – i.e., to ignore IF when making $R_2$ (E. A. Bilodeau, Sulzer, & Levy, 1962; E. A. Bilodeau, Levy, & Sulzer, 1963; E. A. Bilodeau & Levy, 1964).

### *Goal Variability*

The studies collected under this heading give IF as one $f(R)$ on some trials and a different $f(R)$ on others. The transformation may be treated as using different goals on different trials, i.e., $f_1 = (R - G_1)$ vs. $f_2 = (R - G_2)$, etc., or different intercept values, i.e., $f_1 = (R - G)$ vs. $f_2 = (R - G) + X$. There may be only two functions, or several goals or values of $X$. Clearly, studies of misinformative feedback, in which the correct choice is sometimes signaled as incorrect (Morin, 1955), are closely related. Indeed, any treatment in which a given R (S constant) does not receive the same IF on all trials falls reasonably under this section – which could include most of the experiments in probability learning.

The positioning studies treated below fall into three groups: experiments in countertraining in which the subjects first acquire an R of a given amplitude, and then, without instruction, are given $IF_n = (R_n - G) + nk$, on successive countertraining trials; experiments in which the goal changes unsystematically from trial to trial, the subject told that he must guess what goal is in effect; and experiments in which a pair of goals alternates systematically over trial blocks, without the subject's being told that this will be true. The confounding of instructions and variety of goal alternation is not necessary, and Morin's (1955) comparison of instruction treatments makes it clear that outcome is a function of both IF and instructions.

One experiment in countertraining (Ryan & Bilodeau, 1962) has already been reported under Locus (Fig. 5): after eight lever-training trials, 20 units (4°) were subtracted from the obtained discrepancy, $R - 180$, in arbitrary units, on the first trial of countertraining; 40 units on the second countertraining trial; etc. The effect of the first modified IF was to reduce $R_2$ of countertraining ($R_9$ in Fig. 5) by something less than 20 units; further trials continued to reduce R amplitude. Two points, however, should be noted. (a) The subjects did not anticipate the next modification in IF, but corrected for the last change in the error signal. (b) The subjects undercorrected, even when scored for following, rather than anticipating, IF changes. That is, while $IF_n = 180 - 20n$ on the $n$th countertraining trial, $R_n$ approximated $180 - 15(n-1)$, rather than either $180 - 20n$ or $180 - 20(n-1)$.

Tackaberry (1959) used the same apparatus with much the same results in shifting IF by 10 arbitrary units (1.5°) per trial after various numbers of acquisition trials. Overlearned groups (12 or more training trials) decreased R amplitude over fifteen successive countertraining trials. Again, the subjects did not anticipate the next shift in IF in order to receive "a hit"; they did, however, in Tackaberry's study, average a drop of 10 units per trial over the countertraining session.

Rosenberg has used a very attractive apparatus to test properties of his linear, stochastic, feedback model, in both single-subject and dyadic social situations (1962, 1963). The subject turns a knob to control the horizontal position of a pipper on an oscilloscope, while the experimenter or another subject controls vertical input; terminal IF, amount and direction of error, can be read relative to a diagonal line across the face of the scope. Rosenberg has stressed much the same point that is made in continuous IF transformations, and appears so clearly in the countertraining studies—that R change is controlled by the reported difference between required and obtained Rs. Varying the goal from trial to trial according to a normal distribution, Rosenberg (1963) found good agreement between sampling distribution of goal values and distribution of Rs. The average within-subject variance, presented in Fig. 11, shows the effect of the variance of the sampling distribution of goal magnitudes upon the change in R amplitude from trial to trial.

Suppes and Frankmann (1961), testing a different model with another good apparatus, had their subjects guess (R) the goal location on the circumference of a circle of small or large diameter. After he chose, the subject was then shown the goal (required R) for that trial. Again obtained and predicted distributions agreed very closely.

Two other studies in which instructions made no mention of goal shifts (E. A. Bilodeau, 1952c, 1953a) show the effects of systematic, cyclical, rather then random, goal changes, and make, also, a trial-by-trial treatment of R and R change. The apparatus was a lever combined with a display presenting a column of target lights and a matching column of R lights; a single target light was used on all trials. In the first experiment (E. A. Bilodeau, 1952c), subjects alternated over five-trial blocks between treatments in which the

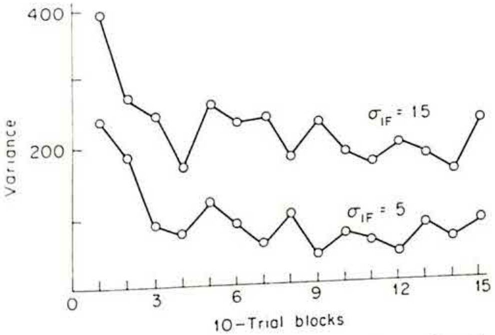


FIG. 11. R variability with variability in the sampling distribution of goal magnitudes. Average within-subject variance over trials. (After Rosenberg, 1963).

R display showed $f_1 = R$ and $f_2 = R + k$; that is, in alternate blocks, a constant was added and the R light signaled a pull longer or shorter than the R signal for the other *f*. Both direction and magnitude of the constant error (or goal shift) proved effective. In unmodified blocks, all groups approached the common lever placement, and in modified blocks they diverged. Undershoot signals produced increased amplitude, in proportion to the size of the error signal; overshoot signals produced proportional decreases in R. The subjects gave very little indication of learning-to-learn; the amount of R shift with shift from one goal to another was more related to $R_n$ and $IF_n$ than to the ordinal number of the cycle. The second experiment (E. A. Bilodeau, 1953a) varied the number of trials in a cycle. Again, with each goal shift, R amplitude changed, but not enough for a hit signal; as trials in the new block continued, obtained Rs more closely approximated the new required R. When the required R changed in a single-trial cycle the subjects adopted a compromise position, tending toward an R halfway between the alternately required Rs, with decreasing between-trial variability over trials. Something of the same sort appeared even in the behavior of groups with 2-, 4-, 6-, and 8-trial cycles; their final R in each cycle showed slight trends (damping) toward the intermediate R over successive cycles. In a two-choice situation, subjects will be influenced by the relative probabilities of the choices' being correct toward a probability-matching solution (Estes, 1964). But when a compromise in R amplitude is possible, it is taken, and it is apparently taken to minimize the frequency of very large discrepancies between required and reported Rs (E. A. Bilodeau, 1953a).

### TRACKING

In tracking tasks scale-grain and full-scale transformations have been studied, as well as interaction with variation in R characteristics. Goal alternation does not make much sense; the required control position varies throughout a trial of either compensatory or following tracking. Though changing the tracking pattern has been studied – for example, by Battig *et al.* (1957) – this does not seem to be as analogous a manipulation as introducing noise into the display (Howell & Briggs, 1959). Some comparable transformations affect tracking much as they affect positioning. Scale-grain transformations combine features of discontinuity and lag, and are detrimental to performance (McConnell & Shelly, 1960). Hunt (1961) has reported data for the functions in Fig. 8, showing that every refinement in the IF reduced tracking error, but with diminishing returns, and, further, with less effect with increasing practice. Small and moderate target-size changes are of little consequence, but large targets do change behavior (Green, 1955); Eason (1963) reports that practice exaggerates the relation between R error and target size. That accuracy may not be affected over a fair range of target sizes is shown in the lack of residuals of previous target size in transfer tests (Gagné & Bilodeau, 1954), though residuals can be produced (Barch, 1953; I. McD. Bilodeau, 1965).

The treatment of display features in continuous responding is as optional as IF in positioning, and variations in tracking displays are more meaningfully treated under the usual labels assigned them in the tracking literature. (The reader should refer to Chapter 8 for a comprehensive treatment of IF transformations in tracking; and to Chapter 9 for a cybernetic treatment of temporal and spatial displacements in sensory feedback.) The most obvious option, and an important one, is the choice of compensatory or following displays (Hartman & Fitts, 1955; Chernikoff, Birmingham, & Taylor, 1955; Gordon, 1959; Obermayer, Swartz, & Muckler, 1962). Others are display-amplification treatments (Battig *et al.*, 1955a; Hartman & Fitts, 1955); spatial displacement (Rhule & Smith, 1959); temporal delay (van Bergeijk & David, 1959); display quickening (Holland & Henson, 1956; F. V. Taylor, 1957); noise (Briggs *et al.*, 1957; Siegel & Robinson, 1963) and the noisy element (Howell & Briggs, 1959); interrupting the target (Battig *et al.*, 1955b); eliminating the follower (Szafran, 1951); and interrupting and eliminating target, follower, or both (Poulton, 1957); various display changes, R demands constant (Fitts & Simon, 1952; Gordon, 1959); and a number of other variations in which display features vary with system dynamics or R requirements (Chernikoff & Taylor, 1957) and in a multitude of control-display variations (Garvey & Knowles, 1954; Hartman & Fitts, 1955; Bennett, 1956; Chernikoff & LeMay, 1963). The number and effectiveness of these treatments is an index of IF's importance in controlling continuous responding.

## References

Adams, J. A. Motor skills. *Annu. Rev. Psychol.*, *1964*, **15**, 181–202.

Alluisi, E. A., & Sidorsky, R. C. The empirical validity of equal discriminability scaling. *J. exp. Psychol.*, 1958, **55**, 86–95.

Ammons, R. B., Ammons, C. H., & Morgan, R. L. Subskills in rotary pursuit as affected by rate and accuracy requirements and by distribution of practice. *J. gen. Psychol.*, 1958, **58**, 259–279.

Anderson, N. H. An evaluation of stimulus sampling theory. In A. W. Melton (Ed.), *Categories of human learning*, New York: Academic Press, 1964. Pp. 129–144.

Andreas, B. G., Finck, A., Green, R. F., Smith, S., & Spragg, S. D. S. Two-dimensional compensatory tracking performance as a function of control-display movement relationships, positioning vs. velocity relationship, and miniature vs. large stick control. *J. Psychol.*, 1959, **48**, 237–246.

Annett, J. Learning a pressure under conditions of immediate and delayed knowledge of results. *Quart. J. exp. Psychol.*, 1959, **11**, 3–15.

Annett, J., & Kay, H. Knowledge of results and skilled performance. *Occup. Psychol.*, 1957, **31**, 69–79.

Archer, E. J., & Namikas, G. A. Pursuit rotor performance as a function of delay of information feedback. *J. exp. Psychol.*, 1958, **56**, 325–327.

Archer, E. J., Kent, G. W., & Mote, F. A. Effect of long-term practice and time-on-target information feedback on a complex tracking task. *J. exp. Psychol.*, 1956, **51**, 103–112.

Atkinson, R. C., & Carterette, E. C. The effect of information feedback upon psychophysical judgments. *Psychon. Sci.*, 1964, **1**, 83–84.

Bahrick, H. P. An analysis of stimulus variables influencing the proprioceptive control of movements. *Psychol. Rev.*, 1957, **64**, 324–328.

Bahrick, H. P., & Shelly, C. Time sharing as an index of automatization. *J. exp. Psychol.*, 1958, **56**, 288–293.

Bahrick, H. P., Noble, M. E., & Fitts, P. M. Extra-task performance as a measure of learning a primary task. *J. exp. Psychol.*, 1954, **48**, 298–302.

Bahrick, H. P., Fitts, P. M., & Schneider, R. Reproduction of simple movements as a function of factors influencing proprioceptive feedback. *J. exp. Psychol.*, 1955, **49**, 445–454.

Barch, A. M. The effect of difficulty of task on proactive facilitation and interference. *J. exp. Psychol.*, 1953, **46**, 37–42.

Battig, W. F., Nagel, E. H., & Brogden, W. J. The effects of error-magnification and marker-size on bidimensional compensatory tracking. *Amer. J. Psychol.*, 1955, **68**, 585–594.

Battig, W. F., Voss, J. F., & Brogden, W. J. Effect of frequency of target intermittence upon tracking. *J. exp. Psychol.*, 1955, **49**, 244–248.

Battig, W. F., Nagel, E. H., Voss, J. F., & Brogden, W. J. Transfer and retention of bidimensional compensatory tracking after extended practice. *Amer. J. Psychol.*, 1957, **70**, 75–80.

Becker, P. W., Mussina, C. M., & Persons, R. W. Intertrial interval, delay of knowledge of results, and motor performance. *Percept. mot. Skills*, 1963, **17**, 559–563.

Bennett, C. A. Sampled-data tracking: Sampling of the operator's output. *J. exp. Psychol.*, 1956, **51**, 429–438.

Bilodeau, E. A. A further study of the effects of target size and goal attainment upon the development of response accuracy. *USAF Hum. Resour. Res. Center res. Bull.*, 1952, No. 52-7. (a)

Bilodeau, E. A. A preliminary study of the effects of reporting goals as a function of different degrees of response accuracy. *USAF Hum. Resour. Res. Center res. Bull.*, 1952, No. 52-4. (b)

Bilodeau, E. A. Some effects of modification of information about a previous response upon the acquisition of two lever positioning habits. *USAF Hum. Resour. Res. Center res. Bull.*, 1952, No. 52-1. (c)

Bilodeau, E. A. Some effects of various degrees of supplemental information given at two levels of practice upon the acquisition of a complex motor skill. *USAF Hum. Resour. Res. Center res. Bull.*, 1952, No. 52-15. (d)

Bilodeau, E. A. Acquisition of two lever-positioning responses practiced over several periods of alternation. *J. exp. Psychol.*, 1953, **46**, 43–49. (a)

Bilodeau, E. A. Speed of acquiring a simple motor response as a function of the systematic transformation of knowledge of results. *Amer. J. Psychol.*, 1953, **66**, 409–420. (b)

Bilodeau, E. A. Accuracy of response as a function of target width. *J. exp. Psychol.*, 1954, **47**, 201–207.

Bilodeau, E. A. Motor performance as affected by magnitude and direction of error contained in knowledge of results. *J. Psychol.*, 1955, **40**, 103–113. (a)

Bilodeau, E. A. Variations in knowledge of component performance and its effects upon part-part and part-whole relations. *J. exp. Psychol.*, 1955, **50**, 215–224. (b)

Bilodeau, E. A. Studies of target size and the control of psychomotor behavior through systematic transformation of knowledge of results. In, G. Finch and R. Cameron (Eds.), *Symposium on Air Force human engineering, personnel and training research.* Washington, D.C.: Natl. Acad. Sci. – Natl. Res. Council (Publ. No. 445), 1956. Pp. 17–24.

Bilodeau, E. A. Inadequate resources, experimental design, and review of the informative feedback literature. *Percept. mot. Skills*, 1965, **20**, 121–122.

Bilodeau, E. A., & Bilodeau, I. McD. Variable frequency of knowledge of results and the learning of a simple skill. *J. exp. Psychol.*, 1958, **55**, 379–383. (a)

Bilodeau, E. A., & Bilodeau, I. McD. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612. (b)

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Bilodeau, E. A., & Levy, C. M. Long-term memory as a function of retention time and other conditions of training and recall. *Psychol. Rev.*, 1964, **71**, 27–41.

Bilodeau, E. A., & Morin, R. E. Proficiency on the Pedestal Sight Manipulation Test with and without the tracking pipper. *USAF Hum. Resour. Res. Center res. Bull.*, 1951, No. 51-27.

Bilodeau, E. A., & Rosenbach, J. H. Acquisition of response proficiency as a function of rounding error in informative feedback. *USAF Hum. Resour. Res. Center res. Bull.*, 1953, No. 53-21.

Bilodeau, E. A., & Ryan, F. J. A test for interaction of delay of knowledge of results and two types of interpolated activity. *J. exp. Psychol.*, 1960, **59**, 414–419.

Bilodeau, E. A., Bilodeau, I. McD., & Schumsky, D. A. Some effects of introducing and withdrawing knowledge of results early and late in practice. *J. exp. Psychol.*, 1959, **58**, 142–144.

Bilodeau, E. A., Sulzer, J. L., & Levy, C. M. Theory and data on the interrelationships of three factors of memory. *Psychol. Monogr.*, 1962, **76**, No. 20 (Whole No. 539).

Bilodeau, E. A., Levy, C. M., & Sulzer, J. L. Long-term retention under conditions of artificially induced recall of related events. *Percept. mot. Skills*, 1963, **16**, 895–910 (Monogr. Suppl. No. 8-V16).

Bilodeau, I. McD. Accuracy of a simple positioning response with variation in the number of trials by which knowledge of results is delayed. *Amer. J. Psychol.*, 1956, **69**, 434–437.

Bilodeau, I. McD. Transfer of training across target sizes. *J. exp. Psychol.*, 1965, **70**, 135–140.

Bilodeau, I. McD., & Rosenquist, H. S. Supplementary feedback in rotary-pursuit tracking. *J. exp. Psychol.*, 1964, **68**, 53–57.

Blick, K. A., & Bilodeau, E. A. Interpolated activity and the learning of a simple skill. *J. exp. Psychol.*, 1963, **65**, 515–519.

Bourne, L. E., Jr. Effects of delay of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201–207.

Bourne, L. E., Jr., & Bunderson, C. V. Effects of delay of informative feedback and length of postfeedback interval on concept identification. *J. exp. Psychol.*, 1963, **65**, 1–5.

Bourne, L. E., Jr., & Pendleton, R. B. Concept identification as a function of completeness and probability of information feedback. *J. exp. Psychol.*, 1958, **56**, 413–420.

Brackbill, Y., & Kappy, M. S. Delay of reinforcement and retention. *J. comp. physiol. Psychol.*, 1962, **55**, 14–18.

Brackbill, Y., Isaacs, R. B., & Smelkinson, N. Delay of reinforcement and the retention of unfamiliar, meaningless material. *Psychol. Rep.*, 1962, **11**, 553–554.

Brackbill, Y., Boblitt, W. E., Davlin, D., & Wagner, J. E. Amplitude of response and the delay-retention effect. *J. exp. Psychol.*, 1963, **66**, 57–64.

Brackbill, Y., Wagner, J. E., & Wilson, D. Feedback delay and the teaching machine. *Psychol. in the Schools*, 1964, **1**, 148–156.

Brackett, H. R., & Battig, W. F. Method of pretraining and knowledge of results in paired-associate learning under conditions of repetition and non-repetition. *Amer. J. Psychol.*, 1963, **76**, 66–73.

Briggs, G. E., Fitts, P. M., & Bahrick, H. P. Learning and performance in a complex tracking task as a function of visual noise. *J. exp. Psychol.*, 1957, **53**, 379–387.

Brown, J. S. A proposed program of research on psychological feedback (knowledge of results) in the performance of psychomotor tasks. In, Research planning conference on perceptual and motor skills, *USAF Hum. Resour. Res. Center conf. Rep.*, 1949, No. 49-2. Pp. 81–87.

Chase, R. A. An information-flow model of the organization of motor activity. U.S. Dept. of Health, Education, and Welfare, Public Health Service, August, 1963.

Chernikoff, R., & LeMay, M. Effect of various display-control configurations on tracking with identical and different coordinate dynamics. *J. exp. Psychol.*, 1963, **66**, 95–99.

Chernikoff, R., & Taylor, F. V. Effects of course frequency and aided time constant on pursuit and compensatory tracking. *J. exp. Psychol.*, 1957, **53**, 285–292.

Chernikoff, R., Birmingham, H. P., & Taylor, F. V. A comparison of pursuit and compensatory tracking under conditions of aiding and no aiding. *J. exp. Psychol.*, 1955, **49**, 55–59.

Chinn, R. McC., & Alluisi, E. A. Effect of three kinds of knowledge-of-results information on three measures of vigilance performance. *Percept. mot. Skills*, 1964, **18**, 901–912.

Conklin, J. E. Effect of control lag on performance in a tracking task. *J. exp. Psychol.*, 1957, **53**, 261–268.

Dees, V., & Grindley, G. C. The effect of knowledge of results on learning and performance. IV. The direction of the error in very simple skills. *Quart. J. exp. Psychol.*, 1951, **3**, 36–42.

Denny, M. R. The role of secondary reinforcement in a partial reinforcement learning situation. *J. exp. Psychol.*, 1946, **36**, 373–389.

Denny, M. R., Allard, M., Hall, E., & Rokeach, M. Supplementary report: Delay of knowledge of results, knowledge of task, and intertrial interval. *J. exp. Psychol.*, 1960, **60**, 327.

Eason, R. G. Relation between effort, tension level, skill, and performance efficiency in a perceptual-motor task. *Percept. mot. Skills*, 1963, **16**, 297–317.

Eaton, M. T. A study of latent learning. *J. exp. Psychol.*, 1935, **18**, 683–707.

Elwell, J. L., & Grindley, G. C. The effect of knowledge of results on learning and performance. I. A co-ordinated movement of the two hands. *Brit. J. Psychol.*, 1938, **29**, 39–54.

Estes, W. K. Probability learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 90–128.

Fitts, P. M. Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 244–285.

Fitts, P. M., & Deininger, R. L. S-R compatibility: Correspondence among paired elements within stimulus and response codes. *J. exp. Psychol.*, 1954, **48**, 483–492.

Fitts, P. M., & Simon, C. W. Some relations between stimulus patterns and performance in a continuous dual pursuit task. *J. exp. Psychol.*, 1952, **43**, 428–436.

Fitts, P. M., Noble, M. E., Bahrick, H. P., & Briggs, G. E. Skilled performance, Part II. *USAF WADC final Rep.*, 1959.

Fleishman, E. A. A relationship between incentive motivation and ability level in psychomotor performance. *J. exp. Psychol.*, 1958, **56**, 78–81.

Foley, P. J., Lavery, J. J., & Abbey, D. S. Scotopic acuity and knowledge of results. *Percept. mot. Skills*, 1964, **18**, 505–508.

Fox, P. W. Learning a discrete motor response with variation in frequency and source of informative feedback. Unpublished master's thesis, Tulane Univer., 1962.

French, E. G. Some characteristics of achievement motivation. *J. exp. Psychol.*, 1955, **50**, 232–236.

Gagné, R. M., & Bilodeau, E. A. The effects of target size variation on skill acquisition. *USAF Personnel Train. Res. Center res. Bull.*, 1954, No. 54-5.

Garvey, W. D., & Knowles, W. B. Response time patterns associated with various display-control relationships. *J. exp. Psychol.*, 1954, **47**, 315–322.

Garvey, W. D., Sweeney, J. S., & Birmingham, H. P. Differential effects of 'display lags' and 'control lags' on the performance of manual tracking systems. *J. exp. Psychol.*, 1958, **56**, 8–10.

Gilbert, R. W., & Crafts, L. W. The effect of signal for error upon maze learning and retention. *J. exp. Psychol.*, 1935, **18**, 121–132.

Goldstein, M., & Rittenhouse, C. H. Knowledge of results in the acquisition and transfer of a gunnery skill. *J. exp. Psychol.*, 1954, **48**, 187–196.

Gordon, N. B. Learning a motor task under varied display conditions. *J. exp. Psychol.*, 1959, **57**, 65–73.

Green, R. F. Transfer of skill on a following tracking task as a function of task difficulty (target size). *J. Psychol.*, 1955, **39**, 355–370.

Greenspoon, J., & Foreman, S. Effect of delay of knowledge of results on learning a motor task. *J. exp. Psychol.*, 1956, **51**, 226–228.

Grice, G. R. An experimental study of the gradient of reinforcement in maze learning. *J. exp. Psychol.*, 1942, **30**, 475–489.

Hamilton, H. C. The effect of incentives on accuracy of discrimination measured on the Galton bar. *Arch. Psychol.*, N.Y., 1929, No. 103.

Hartman, B. O., & Fitts, P. M. Relation of stimulus and response amplitude to tracking performance. *J. exp. Psychol.*, 1955, **49**, 82–92.

Hauty, G. T., & Payne, R. B. Mitigation of work decrement. *J. exp. Psychol.*, 1955, **49**, 60–67.

Holding, D. H. Guidance in pursuit tracking. *J. exp. Psychol.*, 1959, **57**, 362–366.

Holland, J. G., & Henson, J. B. Transfer of training between quickened and unquickened tracking systems. *J. appl. Psychol.*, 1956, **40**, 362–366.

Houston, R. C. The function of knowledge of results in learning a complex motor skill. Unpublished master's thesis, Northwestern Univer., 1947.

Howell, W. C., & Briggs, G. E. The effects of visual noise and locus of perturbation on tracking performance. *J. exp. Psychol.*, 1959, **58**, 166–173.

Hunt, D. P. The effect of the precision of informational feedback on human tracking performance. *Hum. Factors*, 1961, **3**, 77–85.

Karlin, L. Psychological study of motor skills: Phase I. *Nav. Train. Devices Center tech. Rep.*, 1960, No. 558-1.

Kinkade, R. G. A differential influence of augmented feedback on learning and on performance. *USAF WADC tech. Rep.*, 1963, No. 63-12.

Kintsch, W., & McCoy, D. F. Delay of informative feedback in paired-associate learning. *J. exp. Psychol.*, 1964, **68**, 372–375.

Larimer, G. S., & White, J. Some effects of monetary reward and knowledge of results on judgment. *J. exp. Psychol.*, 1964, **67**, 27–32.

Larré, E. E. Interpolated activity before and after knowledge of results. Unpublished doctoral dissertation, Tulane Univer., 1961.

Lavery, J. J. Retention of a skill following training with and without instructions to retain. *Percept. mot. Skills*, 1964, **18**, 275–281.

Lavery, J. J., & Suddon, F. H. Retention of simple motor skills as a function of the number of trials by which KR is delayed. *Percept. mot. Skills*, 1962, **15**, 231–237.

Levy, C. M. Learning of a motor task with action and learning feedback. Unpublished master's thesis, Tulane Univer., 1961.

Lincoln, R. S. Learning a rate of movement. *J. exp. Psychol.*, 1954, **47**, 465–470.

Lincoln, R. S. Learning and retaining a rate of movement with the aid of kinesthetic and verbal cues. *J. exp. Psychol.*, 1956, **51**, 199–204.

Lorge, I., & Thorndike, E. L. The influence of delay in the after-effect of a connection. *J. exp. Psychol.*, 1935, **18**, 186–194.

McConnell, D., & Shelly, M. W. Tracking performance on a sequence of step functions which approaches a continuous function as a limit. *J. exp. Psychol.*, 1960, **59**, 312–320.

McGuigan, F. J. The effect of precision, delay, and schedule of knowledge of results on performance. *J. exp. Psychol.*, 1959, **58**, 79–84.

Mackworth, N. H. Researches on the measurement of human performance. *Med. Res. Council (Brit.), spec. Rep. Ser.*, 1950, No. 268.

MacPherson, S. J., Dees, V., & Grindley, G. C. The effect of knowledge of results on learning and performance. II. Some characteristics of very simple skills. *Quart. J. exp. Psychol.*, 1948, **1**, 68–78.

Melton, A. W. (Ed.), *Categories of human learning*. New York: Academic Press, 1964.

Miller, R. B. Handbook on training and training equipment design. *USAF WADC tech. Rep.*, 1953, No. 53-136.

Morgan, C. L., & Morgan, L. V. Effects of immediate awareness of success and failure upon objective examination scores. *J. exp. Educ.*, 1935, **4**, 63–66.

Morin, R. E. Factors influencing rate and extent of learning in the presence of misinformative feedback. *J. exp. Psychol.*, 1955, **49**, 343–351.

Morin, R. E., & Gagné, R. M. Pedestal Sight Manipulation Test performance as influenced by variation in type and amount of psychological feedback. *USAF Hum. Resour. Res. Center res. Note*, 1951, No. 51-7.

Noble, C. E., & Alcock, W. T. Human delayed-reward learning with different lengths of task. *J. exp. Psychol.*, 1958, **56**, 407–412.

Noble, C. E., & Broussard, I. G. Effects of complex transformations of feedback upon simple instrumental behavior. *J. exp. Psychol.*, 1955, **50**, 381–386.

Noble, C. E., & Noble, J. L. Human trial-and-error learning under joint variation of locus of reward and type of pacing. *J. exp. Psychol.*, 1958, **56**, 103–109.

Noble, C. E., Fuchs, J. E., Robel, D. P., & Chambers, R. W. Individual vs. social performance on two perceptual-motor tasks. *Percept. mot. Skills*, 1958, **8**, 131–134.

Noble, M. E., & Bahrick, H. P. Response generalization as a function of intratask response similarity. *J. exp. Psychol.*, 1956, **51**, 405–412.

Obermayer, R. W., Swartz, W. F., & Muckler, F. A. Interaction of information displays with control system dynamics and course frequency in continuous tracking. *Percept. mot. Skills*, 1962, **15**, 199–215.

Payne, R. B., & Hauty, G. T. Effect of psychological feedback upon work decrement. *J. exp. Psychol.*, 1955, **50**, 343–351.

Pearson, R. G., & Hauty, G. T. Adaptive processes determining proprioceptive perception of verticality. *J. exp. Psychol.*, 1959, **57**, 367–371.

Pearson, R. G., & Hauty, G. T. Role of postural experiences in proprioceptive perception of verticality. *J. exp. Psychol.*, 1960, **59**, 425–428.

Pollack, I., & Johnson, L. Keeping track of the immediately past states of variables. *Percept. mot. Skills*, 1964, **18**, 55–58.

Poulton, E. C. On the stimulus and response in pursuit tracking. *J. exp. Psychol.*, 1957, **53**, 189–194.

Rhule, W., & Smith, K. U. Effects of inversion of the visual field on human motions. *J. exp. Psychol.*, 1959, **57**, 338–343.

Rosenberg, S. Two-person interactions in a continuous-response task. In J. H. Criswell *et al.* (Eds.), *Mathematical methods in small group processes*. Stanford, Calif.: Stanford Univer. Press, 1962. Pp. 282–304.

Rosenberg, S. Behavior in a continuous-response task with noncontingent reinforcement. *J. exp. Psychol.*, 1963, **66**, 168–176.

Ross, S., & Bricker, P. D. The effect of an amount-set on a repetitive motor task. *J. exp. Psychol.*, 1951, **42**, 39–43.

Rouanet, H., & Rosenberg, S. Stochastic models for the response continuum in a determinate situation: Comparison and extensions. *J. math. Psychol.*, 1964, **1**, 215–232.

Ryan, F. J., & Bilodeau, E. A. Countertraining of a simple skill with immediate and 1-week delays of informative feedback. *J. exp. Psychol.*, 1962, **63**, 19–22.

Saltzman, I. J. Delay of reward and human verbal learning. *J. exp. Psychol.*, 1951, **41**, 437–439.

Saltzman, I. J., Kanfer, F. H., & Greenspoon, J. Delay of reward and human motor learning. *Psychol. Rep.*, 1955, **1**, 139–142.

Saufley, W. H., Jr., & Bilodeau, I. McD. Protective self-pacing during learning. *J. exp. Psychol.*, 1963, **66**, 596–600.

Schumsky, D. A. The use of transformations of knowledge of results with negative slope on a simple motor response. Unpublished master's thesis, Tulane Univer., 1959.

Seashore, H., & Bavelas, A. The functioning of knowledge of results in Thorndike's line-drawing experiment. *Psychol. Rev.*, 1941, **48**, 155–164.

Siegel, J., & Robinson, D. N. Prediction of visual tracking proficiency on a A-scope presentation from the measured effects of noise and time. *Percept. mot. Skills*, 1963, **16**, 357–360.

Simon, J. R., & Simon, B. P. Duration of movements in a dial setting task as a function of the precision of manipulation. *J. appl. Psychol.*, 1959, **43**, 389–394.

Smith, K. U. *Delayed sensory feedback and behavior.* Philadelphia, Pa.: W. B. Saunders Co., 1962.

Smith, K. U., Wargo, L., Jones, R., & Smith, W. M. Delayed and space-displaced sensory feedback and learning. *Percept. mot. Skills*, 1963, **16**, 781–796.

Smith, W. M., McCrary, J. W., & Smith, K. U. Delayed visual feedback and behavior. *Science*, 1960, **132**, 1013–1014.

Smith, W. M., Smith, K. U., Stanley, R., & Harley, W. Analysis of performance in televised visual fields: Preliminary report. *Percept. mot. Skills*, 1956, **6**, 195–198.

Smode, A. F. Learning and performance in a tracking task under two levels of achievement information feedback. *J. exp. Psychol.*, 1958, **56**, 297–304.

Solley, C. M. Reduction of error with practice in perception of the postural vertical. *J. exp. Psychol.*, 1956, **52**, 329–333.

Spence, J. T. Verbal discrimination performance under different verbal reinforcement combinations. *J. exp. Psychol.*, 1964, **67**, 195–197.

Spence, K. W. *Behavior theory and conditioning.* New Haven, Conn.: Yale Univer. Press, 1956.

Stockbridge, H. C. W., & Chambers, B. Aiming, transfer of training, and knowledge of results. *J. appl. Psychol.*, 1958, **42**, 148–153.

Sulzer, J. L. Arc versus place skills after guided and knowledge of results training. Unpublished master's thesis, Tulane Univer., 1961.

Suppes, P. Stimulus-sampling theory for a continuum of responses. In K. J. Arrow *et al.* (Eds.), *Mathematical methods in the social sciences, 1959.* Stanford, Calif.: Stanford, Univer. Press, 1960. Pp. 348–365.

Suppes, P., & Frankmann, R. W. Test of stimulus sampling theory for a continuum of responses with unimodal noncontingent determinate reinforcement. *J. exp. Psychol.*, 1961, **61**, 122–132.

Szafran, J. Changes with age and with exclusion of vision in performance at an aiming task. *Quart. J. exp. Psychol.*, 1951, **3**, 111–118.

Tackaberry, T. H. Counterconditioning in human beings as a function of degree of overtraining. Unpublished master's thesis, Tulane Univer., 1959.

Taubman, R. E. The effect of practice with and without reinforcement on the judgment of auditory number. *J. exp. Psychol.*, 1944, **34**, 143–151.

Taylor, A., & Noble, C. E. Acquisition and extinction phenomena in human trial-and-error learning under different schedules of reinforcing feedback. *Percept. mot. Skills*, 1962, **15**, 31–44.

Taylor, F. V. Simplifying the controller's task through display quickening. *Occup. Psychol.*, 1957, **31**, 120–125.

Taylor, F. V., & Birminghan, H. P. Studies of tracking behavior. II. The acceleration pattern of quick manual corrective responses. *J. exp. Psychol.*, 1948, **38**, 783–795.

Thorndike, E. L. *Human learning*. New York: Century, 1931.

Trowbridge, M. H., & Cason, H. An experimental study of Thorndike's theory of learning. *J. gen. Psychol.*, 1932, **7**, 245–260.

Underwood, B. J. *Experimental psychology*. New York: Appleton-Century-Crofts, Inc., 1949. Pp. 414–416.

van Bergeijk, W. A., & David, E. E., Jr. Delayed handwriting. *Percept. mot. Skills*, 1959, **9**, 347–357.

Walton, D., & Begg, T. L. The effects of incentives on the performance of defective imbeciles. *Brit. J. Psychol.*, 1958, **49**, 49–55.

Warrick, M. J. Effect of transmission-type control lags on tracking accuracy. *USAF WADC tech. Rep.*, 1949, No. 4916.

Weinberg, D. R., Guy, D. E., & Tupper, R. W. Variation of postfeedback interval in simple motor learning. *J. exp. Psychol.*, 1964, **67**, 98–99.

Williams, A. C., & Briggs, G. E. On-target versus off-target information and the acquisition of tracking skill. *J. exp. Psychol.*, 1962, **64**, 519–525.

# Information Feedback

## Comments on Professor I. McD. Bilodeau's Paper

L. E. BOURNE, Jr.[1]
*University of Colorado*

Dr. Ina Bilodeau has provided an extensive review of recent literature pertaining to the effects of variations in informative feedback (IF) on motor-skills learning. This work contains not only a fair appraisal of our current knowledge but also a sharp focus on some empirical and semitheoretical issues which remain to be resolved. The chapter was written, clearly, by an expert. Any significant gaps in its coverage, interpretive distortions in its presentation of data, and the like, if indeed they exist, are not obvious to me. Thus, I find the role of critic, taken in the usual sense, a difficult one to fulfill.

For this reason, coupled with the fact that my own expertise in the general area of motor learning is next to negligible, I address these remarks primarily to questions concerning the generality of conclusions drawn by Dr. Bilodeau. In essence I am accepting her invitation (p. 259) to fill in some of the holes. It is, of course, patently impossible to review the remaining literature on IF in a few short pages and at a depth comparable with that of Bilodeau's summary. My goal is much more modest. What I hope to indicate, by citing a few current studies, are cases where manipulations of the same variable have similar or different effects on behavior in learning situations usually considered to be outside the area of skills. While my effort may amount to little more than an intellectual exercise, it embodies at least the possibilities of suggesting (a) important continuities across various categories of learning and (b) alternative interpretations of the influence of IF on human behavior.

---

[1] Preparation of the manuscript and some of the research reported herein were supported in part by grant MH-08315 from the National Institute of Mental Health, U. S. Public Health Service. I express deep appreciation to Professor W. F. Battig who, in my absence, presented these comments at the Conference.

## Some General Issues

A few problems need to be considered—though I hold out no great hope of resolving them—before we turn to the available data. First of all, there are clear differences in performance requirements between the skills tasks reviewed by Bilodeau and many of the human-learning tasks which I shall discuss. By definition, the emphasis in skills research is on some facet of response learning—describable either as an integration of elemental acts into some larger functional unit or as the shaping of a rough response with respect to an experimenter-defined standard. In other learning situations, the response or responses to be made often already exist in sufficiently well-learned form within the subject's behavioral repertoire; the response (or responses) is available and the task requires no significant change in its structure. Second, it is typical of skills experiments to explore the changes in characteristics of a single response, if we may so define the act of positioning a lever or tracking a target. Subjects usually are not burdened with the necessity to acquire several unique, differentiable behaviors within a single task. On the other hand, this in a sense is exactly the process explored in many nonskills experiments. Third, in tasks designed for the study of skills, there usually exist no experimenter-governed stimuli or cues, which elicit or in any sense determine the correct response, prior to response-execution. The association of response (or responses) with discriminable stimulus (or stimuli), which is the event of fundamental interest in many learning situations, is generally not an issue in the acquisition of a motor skill.

Each of the aforementioned distinctions among learning tasks and learning experiments is admittedly artificial and arbitrary. Each can legitimately be challenged. It would be a mistake to assert that response learning is studied only by the skills researcher. As Underwood and others (Underwood & Schulz, 1960) have so convincingly demonstrated, the response-integration (e.g., learning to pronounce) or response-recall phases of verbal paired-associates learning may be appreciable when unfamiliar, hard-to-pronounce, or highly similar responses are a part of the task. While it is typical of the skills experimenter to study the growth of single habits, surely there is nothing to prevent an expansion of tasks to incorporate multiple habits, which is probably more characteristic of everyday skilled performance anyway. Finally, the distinction between associative and response learning processes seems, on close scrutiny, to be merely a terminological convenience, which permits the separation of tasks wherein external stimuli controlled by the experimenter are or are not present as guides to proper responding. While we usually think of the associative stage of learning as a matter of "hooking-up" Rs with Ss, who is to deny that the same process is involved in at least many facets of response learning. It seems quite likely, for example, that learning and perfecting the so-called continuous act of tracking is essentially an associative process wherein the stimuli exist in the perceptible external or internal environment of the subject, though under somewhat reduced experimenter control.

Thus the task distinctions pointed to in these prefatory remarks are fictions. Each in reality arises from a continuum along which it is impossible to draw any sharp lines. Still they are convenient fictions for they help in this attempt to understand why the effects of a given manipulation of IF may or may not be the same in different experiments. Bilodeau has shown how these variables influence performance in tasks which require the shaping or integration of single responses. We shall consider how they affect behavior in multiresponse tasks, wherein the responses are more or less well learned to begin with and the critical learning resides in unique associations between these responses and experimenter-controlled external stimuli.

## The Function of IF

*Informative feedback* is a term which refers to some signal occurring after or at the time of response, providing an indication of the correctness, accuracy, or adequacy of that response. The information contained in the signal is necessary for learning and probably also for the maintenance of the response. As Bilodeau noted the provision of IF is often intrinsic to a motor task; for example, the subject can determine response accuracy in many tracking tasks simply by paying attention to the target and his own behavior. For proper experimental study, however, IF should be extrinsic and controllable by the experimenter, and Bilodeau was at some pains to show how this control can be effected.

In a positioning task, IF provides not only a necessary but also probably a sufficient amount of information to learn. Or to hedge a bit, whatever else the subject must know, say, a "feel" for location or distance, resides within him. In other learning tasks, where alternative responses are possible and associative processes are more clearly evident, the necessary and sufficient information is spread across two sources, the stimulus, i.e., the preresponse signal, and IF, i.e., the postresponse signal, both of which may be regulated by the experimenter. In any paired-associates task, of course, the subject is at a loss to learn lacking either the stimulus or the IF signal, even if the required response is well known. In such tasks, we may say that the preliminary stimulus determines the response to be made on any particular trial and the IF signal indicates the response to be made on some later trial(s).

This much seems indisputable. A question arises, however, about the possibility of additional consequences of IF. Bilodeau has drawn the issues clearly: Is it the case that reinforcing (e.g., rewarding and punishing) and motivating (e.g., activating and arousing) properties are also attributable to IF? Any answer to the question will be difficult to defend. The action of events which follow a response upon the later behavior of the same organism is one of psychology's persistent problems. My own bias is to stress the guiding or informational function of postresponse events. It is difficult for me to imagine a drive to which "Right" or "Wrong" spoken by an experimenter would appeal, though I am aware of current assertions about a need for approval. Further it would seem that a subject carries with him a level of motivation or arousal, perhaps supplemented by preliminary "pep-talk" instructions, sufficient to the task and that bland periodic reports of response accuracy provided by the experimenter produce little or no appreciable change in this state of affairs. In most circumstances, then, it seems superfluous to attribute reinforcing or motivating functions to IF.

But my comments merely express an opinion and are not meant to deny that, under certain conditions, IF may in fact carry reinforcing or motivating components. Hillix and Marx (1960), for example, were able to show some slight effect of "Right" signals following correct responses in multiple choice learning tasks over and above that which could be accounted for by their sheer informational content. It is, however, interesting to note that the beneficial effect of this "reward" appeared in performance on a transfer task; in original learning, performance was actually affected adversely. In another study, Wilkinson (1961) showed that the deleterious effect of sleep loss on performance in a multiresponse reaction time task can be overcome by supplying augmented IF and interpreted the result in terms of the arousal or motivational properties of IF. These studies, however, provide little reason to alter my original position. The reinforcing function of typical IF is at best minimal and even may have consequences opposite to those usually supposed. The motivating function appears to be important only in those extreme circumstances which conduce to a low level of "natural" arousal. And, in any case, the informational content of IF signals appears crucial to learning and performance.

One final issue bears brief consideration. I have asserted that IF signals carry potential or usable information. It seems reasonable to ask what the nature of that information is and how it operates to affect behavior. This is another complicated question for which no single answer suffices in all cases. Two points are relevant, however. The first is contained in a distinction drawn by Annett and Kay (1957) between knowledge of results and knowledge of the correct response. Typically, the postresponse signal supplied by an experimenter indexes (a) something about the response that should have been made and (b) the departure of the actual response from it. Are both kinds of information necessary to learning? I suspect the answer depends on the extent to which there exist stimuli intrinsic to the task which themselves provide type b information. But the correctness of this assumption is difficult to assess for lack of convenient experimental techniques for separating the types of information. Studies of guided practice hold some promise in this respect, however, and I am impressed though not totally convinced by the results of Holding (1959), who compared the acquisition of pursuit tracking performance in groups which were given normal and guided practice on several preliminary practice trials. Holding reports that full guidance, with the subject observing the target and holding the mechanically moving control, produced as much learning as normal practice. Because the automatic tracking device was perfect, no error (departure from correct response) was fed back to the subject during guided trials. Rather the feedback (if we may call it that) during guidance contained information only about the correct response. The fact that performance on later trials was the same for guidance and normal practice groups led the author to conclude that knowledge of correct response is necessary for skills acquisition and that knowledge of results (or error) is merely incidental. The generality of this conclusion may be questioned on the grounds that (a) only 2 minutes of preliminary guidance or normal practice was provided and (b) the extent to which intrinsic error information is provided by the task is indeterminate. However, results do seem to point up the sufficiency of information about the correct response under some circumstances.

The second point relates to the cite of action of IF on the behavioral chain which constitutes performance on any trial in a learning task. It is clear that IF has the three empirical properties listed by Bilodeau, viz., R strengthening, R sustaining, and error eliminating. But how are these effects accomplished? This question, like others we have raised, is the subject of considerable current debate. There are those who would argue forcefully that IF operates to establish and strengthen associations between (external or response-produced internal) stimuli and overt responses. But it has become popular in recent times to speak of response or associative strengthening as an indirect consequence of the effect of IF on some expectancy, hypothesis, or mediating stimulus-response connection established within a subject. There is a degree of reasonableness in both positions. When the response demanded by the task does not exist or exists only in unimproved form in the subject's repertoire of behaviors, it is difficult to imagine the formation of any relevant hypotheses about action. On the other hand, there is, as we've noted, a large class of learning situations for which the responses exist in highly integrated form at the outset and in which the goal is to learn to make a response "on cue." Here a convincing case can be developed for the importance of mediational events or internal hypotheses. The conclusion that seems inescapable, if the preceding hunches have any validity, is that the development of hypotheses is somehow tied to or a product of the acquisition of responses (and, of course, the acquisition of associations). This, as I understand it, is essentially the position taken recently by Mandler (1962) when he suggested the possible establishment within a subject of symbolic analogs of overt behavior. While IF and/or reinforcement may act initially to promote associative response strength, they operate in later situations, wherein the formerly learned response may or may not be appropriate, on the corresponding analog or hypothesis. Responding then comes to resemble an all-or-none by-product of hypothesis selection and test. The test information is, of course, provided by IF.

We proceed now to a consideration of the effects of variations in IF in multiresponse tasks, where the responses are at least in large measure already available, and in which the emphasis is upon associating these responses with experimenter-regulated stimuli.

## Temporal Locus

The effects of variation in the timing of critical intratrial events, i.e., stimulus, response, and IF, are remarkably consistent across experiments using a variety of learning tasks. Bilodeau has reviewed the evidence from positioning and some response-selection studies which point to the importance of length of interresponse interval and cast serious doubt on the efficacy of empty (unfilled) delay between response and IF. In so doing, she dismissed simple trace theories of stimulus (and response) decay and concluded, cogently, that important processes bound up in performance may take place before, during, and after the IF signal of any trial. She has pointed quite reasonably to interpolated activities required by the task or engaged in spontaneously by the subject as a critical source of interfering (or possibly facilitating) intratrial effects.

### CONCEPT IDENTIFICATION

The influence of short, unfilled delays of IF appears to be negligible both in simple skills learning and in some more complicated situations. It has already been noted that the degrading effect on concept identification—a task which requires the formation of associations between a set of available responses and certain features of multidimensional stimuli—originally assigned to IF delay by Bourne (1957) is an artifact of a failure to control length of post-IF interval. In a study in which these factors were varied orthogonally (Bourne & Bunderson, 1963), only length of post-IF interval produced a reliable change in performance. The latter effect, interestingly, was facilitative, which led the authors to reject interpretations based on the immediate and automatic strengthening property of IF on S-R associations and to speculate about possible conditioning or information-processing operations which may persist into the post-IF interval. It may be recalled that length of interresponse interval (delay plus post-IF) had adverse effects in some positioning studies reported by Bilodeau, which suggests possible inconsistencies in results across experiments. One source of the inconsistency, however, seems reasonably well identified in the work of Weinberg, Guy, and Tupper (1964). These experimenters noted that the above-mentioned investigations dealt with nonoverlapping ranges of post-IF intervals. Durations used in the concept identification studies were less than 9 seconds while in motor-skills experiments, the same intervals were never less than 8 seconds. The data of Weinberg *et al.* helped to understand and to unify experimental results by showing that an increase of short post-IF intervals from 1 to 5 seconds facilitates simple motor learning, as well as concept identification. They concluded that, as the length of this interval increases, so does the subject's opportunity to make use of information provided on any trial. Further, they argued that intratrial learning in simple motor tasks is effectively complete after some short time (approximately 5 seconds) and that, with longer intervals such as those used by Bilodeau and Bilodeau (1958), forgetting may occur. The data imply the existence of an optimal post-IF interval for any task, as the combined result of simultaneous acquisition and forgetting processes.

Following up this lead, Bourne, Guy, Dodd, and Justesen (1965) undertook two experiments to explore a broader range of post-IF intervals in a concept identification task. In the first, they were able to demonstrate an optimal interval which, as is shown in Fig. 1, differed in value with task complexity. For simpler problems, involving the identification of two relevant from among three variable stimulus dimensions, the optimal interval was approximately 9 seconds, but for more complex tasks, with two relevant and five irrelevant dimensions, the optimum was about 17 seconds. The second study made it likely that the degrading effect of longer intervals can be attributed, in part, to forgetting. In it, performance under the normal condition, in which neither the stimulus nor the IF signal remains for inspection during the post-IF interval, was compared with that obtaining for cases in which either the stimulus or the IF signal or both was made available throughout each trial. While availability of IF alone had little beneficial influence (suggesting the general ineffectiveness of duration of IF in these problems), availability of

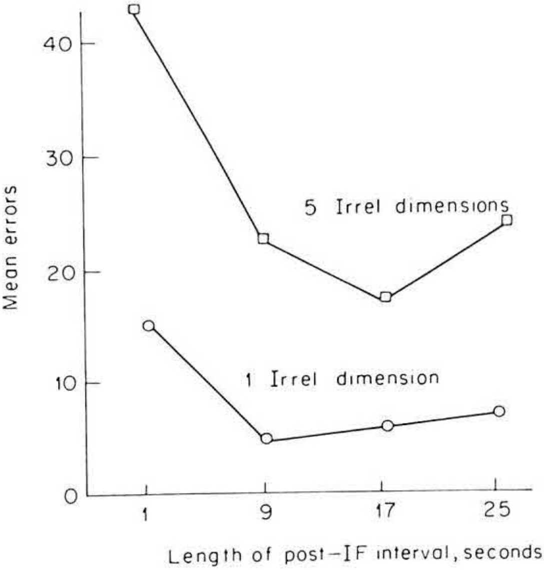


FIG. 1. Mean errors to solution of four-response concept identification problems as a joint function of problem difficulty, i.e., number of irrelevant dimensions, and length of postfeedback interval. Data from Bourne, Guy, Dodd, and Justesen (1965), Exp. I.

the stimulus or (especially) the stimulus and feedback prevented the usual deterioration of performance with long post-IF intervals.

### VERBAL LEARNING

In concept identification, the subject learns a scheme for categorizing geometric designs. The category responses are button presses—simple, well-learned motor acts. The effects of manipulating the time characteristics of critical intratrial events in these problems are quite consistent, though perhaps more complicated, with analogous variations in motor-learning tasks, despite the obvious, large task differences. In light of this evidence, it seems worthwhile to press on, asking whether similarities hold under other changes of task features. There are a few studies of verbal learning in which temporal locus of IF has been varied. Jones and Bourne (1964) reported five relevant experiments. The first two were attempts to replicate the results of Saltzman (1951), which (as Bilodeau has pointed out) are among the few providing strong evidence of an adverse effect of delayed IF on human performance, here in the verbal maze. Saltzman's procedures were followed as closely as the published description permits. The subject's task was to learn which member of each of six pairs of four-digit numbers was correct and to indicate its location, right or left, on each presentation. The two replications differed from each other only in that greater emphasis was placed on memorizing the correct numbers in the second. In the immediate IF condition, each pair was displayed for 6 seconds, after which the correct answer was presented for an additional 6 seconds. For delay, another 6-second unfilled interval was introduced between stimulus and IF. The results of these replications were virtually identical, but inconsistent with Saltzman's findings. Errors and trials to criterion were about the same for immediate and delayed IF. It is, of course, impossible to rule out completely instructional and other procedural factors as sources of the discrepancy between these findings and those of Saltzman. For example, Saltzman reported that his subjects engaged in a noticeable amount of overt rehearsal of the number pairs during unfilled intervals whereas there was no evidence of similar activity in the Jones–Bourne experiments. If subjects are convinced that memorizing the numbers is crucial and concentrate on this task, rehearsal during delays could interfere with retention of choice responses and hence result in an over-all performance decrement. In this case, of course, rehearsal of irrelevant responses, not delay of IF per se, is the variable affecting performance.

The third experiment, employing the traditional paired-associates paradigm, was patterned after that of Bourne and Bunderson (1963). Length of delay between stimulus and IF – 0, 3, and 6 seconds – was varied orthogonally with length of the post-IF interval – 0, 3, and 6 seconds. Both stimuli and responses were difficult-to-pronounce, low-frequency trigrams (Underwood & Schulz, 1960). The time intervals were unfilled; the subject responded by spelling the response term; and IF appeared in the form of the correct S-R pairing, henceforth called the S-R mode. Lengthening both intervals improved performance to essentially the same reliable degree.

These results, taken in conjunction with those of Saltzman (and the replications), suggest that there are at least three sets of variables underlying the effects of delayed IF, which have neither been isolated nor systematically manipulated in studies of human multiple-response learning. One set includes the amount and kind of activity the subject engages in during the delay, a point which has been made by Bilodeau. The second relates to the form or mode of IF. Strictly speaking, if making an overt response to each stimulus is a necessary feature for learning, then any unambiguous information on the correctness of that response should be maximally effective. Intuitively, however, there seems to be considerably more useful information toward the formation of correct associations when IF is provided in the S-R mode than when only the correct response – the R-only mode – is used. For one thing, the S-R mode satisfies a basic condition for associative learning by providing contiguity between stimulus and response members (Spear, Ekstrand, & Underwood, 1964). Finally, there is the question, of some pertinence to skills research, about the possibility of facilitative effects of IF delay on learning in multiple-response tasks wherein the responses are difficult, at least at the outset.

Jones and Bourne reported two further experiments which provide evidence on the importance of the first two variables. In the first, the effects of IF delay were explored under circumstances in which the subject engaged in task-relevant activity during the interval. The trials-delay technique described by Bilodeau, which required the subject to respond to other stimuli while awaiting feedback on an earlier S-R pair, was used. This procedure has important advantages for the study of delay. First, total time per trial is the same regardless of delay length and, second, there is no difference in total activity per trial, thus obviating the necessity to impose varying amounts of irrelevant responding on subjects. The learning materials were hard-to-pronounce trigrams and IF appeared in the S-R mode. The results showed a mild inhibitory effect of number of responses interpolating between any given stimulus (and the subject's response to it) and the appropriate IF signal. Noteworthy is the fact that in the extreme, the trials-delay technique results in the complete separation of learning (observing a correct S-R pair) and performance (responding to a stimulus) properties of the paired-associates task. Battig and Brackett (1961) have found this an easier condition for list-learning than the more typical alternation of learning and performance functions. While the difference was statistically unreliable in the present experiment, performance in the extreme trials-delay condition was best.

The final experiment in this series was a replication of the preceding, except that IF appeared in the R-only mode, i.e., the stimulus was not reinstated by IF. Here the effects were quite dramatic. Performance became increasingly worse with length of trials delay.

One additional study undertaken by Jones (1963) fits nicely with the preceding results. It had the primary purpose of exploring possible interactive effects of delay and IF mode in verbal paired-associates learning but also involved variations in the length of post-IF interval and in the nature of the subject's activity during delay. To summarize briefly, increasing the post-IF interval from 0 to 6 seconds improved performance. The delay effect varied, however, as a function of IF mode and of interpolated activity. Empty 6-second delay facilitated performance with the S-R mode but was no different from the no-delay condition with the R-only mode. Performance was virtually the same with immediate IF for the two modes, but the R-only mode was more difficult when IF was delayed. Counting backward during IF delay inhibited performance, especially when IF supplied only the correct response without reinstatement of the stimulus.

As a whole, these experiments indicate that, in verbal paired-associates learning, with difficult responses, delay may facilitate or interfere with performance depending on (a) the form or mode of IF, and (b) the amount and kind of activity the subject engages in during the delay. Facilitation results under conditions wherein lengthening the delay (and/or post-IF interval) contributes solely or primarily to rest or distribution of practice. Although manipulations of intratrial rests are almost unique to these studies, there is no obvious reason to believe that they operate any differently from intertrial intervals. Underwood (1961) has attributed the effects of intertrial distribution to the opportunity it provides for repeated recovery and extinction of interfering responses. The longer the intertrial interval, the greater the net extinction of interfering responses, yielding superior overall performance. It is clear that increasing time within the list (delay and post-IF intervals) allows for spontaneous recovery of competing responses in much the same way as increasing the time between list repetitions and, as we have seen, may also produce marked facilitative effects. The effect identified with the post-IF interval is, of course, essentially the same as that reported by Bourne and Bunderson (1963) and by Weinberg *et al.* (1964). Although the underlying processes may be different in motor skills, concept identification, and verbal learning, being necessarily associated in part with the performance requirements of these tasks, available evidence now suggests a surprising degree of continuity across human-learning situations.

We may note in passing that these results also are not inconsistent with recent studies of short-term memory, a matter which we shall learn more about in the next chapter. In general, the introduction of an IF delay necessitates memory of either the stimulus or the subject's response to it, or both. The probability of forming a correct association between stimulus and response members will be a direct function of the probability of retaining the necessary information through the delay. The procedure of the trials-delay experiment with R-only IF is particularly noteworthy, for it confounds three variables which have known deleterious effects on short-term memory. Specifically, with respect to the retention of any particular stimulus, increasing delay produces concomitant increases in (a) opportunity for retroactive interference from intervening stimuli, responses, and IF signals (Murdock, 1963); (b) opportunity for proactive interference from previously presented items (Keppel & Underwood, 1962); and (c) number of responses to be recalled during any test sequence (Lloyd, Reid, & Feallock, 1960). While the independent effects of these factors in the data are not assessable, it is to be expected that the ease of associative hookup of S-R pairs and therefore final mastery of a list will vary with each. The general outcome supports the contention of Peterson and Peterson (1959) that overall performance in a verbal learning task is a function of variables affecting the short-term retention of individual items.

## Other Variables

I have devoted more than a proportional amount of time to temporal locus of IF. The remaining variables discussed by Dr. Bilodeau, though no less important in their effects on complex behavior, can be treated only briefly. My plan is to center these remaining comments on examples from the experimental literature, which of course precludes complete coverage or the justification of broad generalizations, but has the compensating advantage of specificity and precision.

### FREQUENCY OF IF

There appear to be very few studies of IF frequency outside the areas of animal learning, conditioning, and motor skills. There are no investigations, analogous to Bilodeau, Bilodeau, and Schumsky (1959), for example, of the effects of withdrawing of IF at some point in the trials sequence during multiresponse associative learning. Indeed, such a study would be of doubtful interest when, as is typical of these tasks, responses need not be created but are simple, well-integrated, and available at the outset. While Bilodeau suggests the possible relevance of probability learning, this work really appears to be more appropriately placed under the heading of IF transformations, for characteristically, IF (though capricious) is available on all trials in such a task.

One type of variation in frequency is of some interest, in light of remarks on possible reinforcing effects of IF. The problem arises from a continuing concern among researchers about the influence of confirming and infirming feedback in associative learning tasks. Buss and Buss (1956) among others have compared the efficacy of the following three feedback conditions. (1) R-W: The experimenter says "Right" after every correct response made by the subject and "Wrong" after every incorrect response. (2) R-N: The experimenter says "Right" after correct responses but nothing after errors. (3) N-W: The experimenter says "Wrong" after errors but nothing after correct responses. Typically, it has been found that performance, measured in trials or errors to criterion, is significantly worse with R-N than with R-W and N-W, the latter conditions not differing. Reinforcement theory has a difficult time with these data unless one agrees with Buss and Buss that "Right" is a much weaker positive reinforcer (reward) than "Wrong" is a negative reinforcer (punishment). This interpretation assumes that learning in card-sorting tasks is predicated on the formation of S-R bonds. An equally plausible account of the data can be developed (Spence, Lair, & Goodstein, 1963) on an assumption that subjects test hypotheses in order to arrive at solution. According to this position, R-W and N-W lead to faster learning because "Wrong" is a less ambiguous form of feedback than "Right." "Wrong" informs the subject that his hypothesis is incorrect and leads him to search for and select another. "Right" may be misleading, for it is possible to make a correct response on the basis of an incorrect hypothesis.

One critical feature, often overlooked, which distinguishes among R-W, R-N, and N-W conditions is frequency of IF. It is common in these experiments to use tasks in which the number of possible responses is greater than two. For example, Buss used the four-category Wisconsin Card Sorting Test (WCST). Because by chance errors are more probable than correct responses in these tasks, more trials accompanied by IF are given to subjects in N-W than in R-N condition. Even if "Right" and "Wrong" are equally reinforcing (though of opposing value) or equally informative (which would be my own preferred way of looking at the problem), N-W may be superior to R-N for reasons of IF frequency. To explore the question, I had two students (D. Guy and N. Wadsworth) collect data from subjects in two successive WCST problems under the usual R-W, R-N, and N-W conditions plus several others as displayed in Fig. 2. The findings were quite clear; trials to criterion were fewer in the R-W and N-W groups than in the R-N group. Performance according to this measure got worse as the percentage of error trials and the percentage of correct response trials accompanied by IF decreased, though the percentage of error trials was the more influential variable. However, when we counted merely the total number of feedback trials to criterion, there were no significant differences over the whole experiment. The impression left by these data is that possible differential reinforcing values or ambiguities in "Right" and "Wrong" IF are negligible relative to frequency of IF. These data seem clearly to indicate that solving WCST problems takes, on the average, a certain number of informed trials, regardless of the correctness of the subject's response.

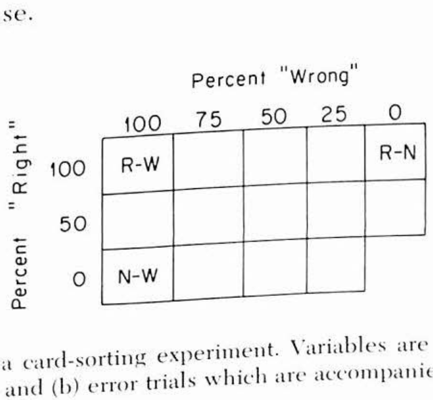


FIG. 2. Design of a card-sorting experiment. Variables are the percentages of (a) correct response trials and (b) error trials which are accompanied by feedback.

### AUGMENTED IF

I gather from Bilodeau's account that augmented IF is effective in improving skills performance only to the extent that it adds usable information to standard (presumably improverished) IF. In contrast to our knowledge about other problems of current concern, this seems to be a quite logical and reasonable state of affairs. I like to think that the same conclusion holds for nonskills learning as well, and we have some evidence to that effect in the Bourne and Pendleton (1958) experiment which shows that concept identification can be expedited by providing more complete IF on the nature of the subject's errors.

One might liberalize this topic to allow the question: "What happens when rewarding or punishing events are combined with neutral IF signals?" While we have by no means settled the issue of possible reinforcement effects inherent in any IF signal, it is surely possible to find reinforcers which can also serve an information-giving function. There is a considerable literature on the effects of monetary and other rewards on human learning (e.g., Miller & Estes, 1961), though rather few comparing IF in this form to neutral IF. One study by Wallace (1964) which makes such a comparison is worth mentioning. Wallace used a concept formation problem in which verbalized guesses or hypotheses about the solution were periodically elicited from subjects. For different groups, a neutral signal (sound) or a derogating remark from the experimenter accompanied all incorrect hypotheses. The effects were complex, but we need simply note that subjects who were punished by the experimenter's verbalized statements offered fewer hypotheses and attained solution more quickly than those receiving the nonverbal signal. While the two types of feedback contained, as far as can be determined, the same information with respect to the subject's behavior, they produced reliably different performance levels. As such the data attest not only to the important effects of reinforcement in complex problems but also (conceivably) to certain beneficial consequences of IF augmented by reinforcement.

### TRANSFORMATIONS OF IF

It is difficult to think of good examples from nonskills research of the equivalent of scale-grain or scale-distortion variables discussed by Dr. Bilodeau. However, there is a fair amount of work pertinent to the issue of goal variability, including the studies of solution shifts in discrimination learning and simple concept formation (Kendler & Kendler, 1962) and the monumental work on probability learning. I am particularly intrigued with the observation of E. A. Bilodeau (1953, reported on p. 287 of the preceding chapter) of a tendency toward compromise in response amplitude in a positioning task when either of two amplitudes could be correct on any given trial. One might wonder whether the compromise, presumably adopted to minimize the frequency of very large discrepancies from IF, might not be affected by the probability of the two possible IF signals, as an analog of the probability-matching solution might suggest. That is, will the actual response given by subjects tend toward an amplitude which is predictable from knowledge of IF probabilities, being closer to the more frequent signal? I have no ready answer to this question for I am unaware of any pertinent data. But while the mood for raising questions exists, it may also be worth asking about the permanency of the subject's adopted compromise in response amplitude. In simple concept identification problems, modified to permit the introduction of a certain percentage of misinformative trials, i.e., trials on which correct signals sometimes accompany incorrect responses and vice versa, subjects are known to give up any attempt to out-guess the experimenter and to make consistently (maximize?) the response more frequently associated with confirming IF (Bourne, 1963).

## Concluding Remarks

Dr. Bilodeau's review and evaluation of the role of IF in motor-skills learning serves the important function of clarifying certain empirical and theoretical issues. She has provided a firm base from which researchers interested in this fertile area may work to elaborate further our understanding of the many problems that remain. I have tried to show, mostly through examples, that many of the same questions may be asked in situations which emphasize facets of behavior other than skills acquisition. There is, I think, surprising continuity in effects of IF variables across tasks. If this is true, we have much to learn about learning in general from skills research. But our knowledge, though considerable, is far from complete and I trust that many of us will continue to concern ourselves with IF variables not only from a variety of theoretical viewpoints but also in a variety of learning tasks.

## REFERENCES

Annett, J., & Kay, H. Knowledge of results and 'skilled performance.' *Occup. Psychol.*, 1957, **31**, 69–79.

Battig, W. F., & Brackett, H. R. Comparison of anticipation and recall methods in paired-associates learning. *Psychol. Rep.*, 1961, **9**, 59–65.

Bilodeau, E. A. Acquisition of two lever-positioning responses practiced over several periods of alternation. *J. exp. Psychol.*, 1953, **46**, 43–49.

Bilodeau, E. A., & Bilodeau, I. McD. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612.

Bilodeau, E. A., Bilodeau, I. McD., & Schumsky, D. A. Some effects of introducing and and withdrawing knowledge of results early and late in practice. *J. exp. Psychol.*, 1959, **58**, 142–144.

Bourne, L. E., Jr. Effects of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201–207.

Bourne, L. E., Jr. Long-term effects of misinformative feedback upon concept identification. *J. exp. Psychol.*, 1963, **65**, 139–147.

Bourne, L. E., Jr. & Bunderson, C. V. Effects of delay of informative feedback and length of postfeedback interval on concept identification. *J. exp. Psychol.*, 1963, **65**, 1–5.

Bourne, L. E., Jr., & Pendleton, R. B. Concept identification as a function of completeness and probability of information feedback. *J. exp. Psychol.*, 1958, **56**, 413–420.

Bourne, L. E., Jr., Guy, D. E., Dodd, D., & Justesen, D. R. Concept identification: the effects of varying length and informational components of the intertrial interval. *J. exp. Psychol.*, 1965, **69**, 624–629.

Buss, A. H., & Buss, E. H. The effect of verbal reinforcement combinations on conceptual learning. *J. exp. Psychol.*, 1956, **52**, 273–278.

Hillix, W. A., & Marx, M. H. Response strengthening by information and effect in human learning. *J. exp. Psychol.*, 1960, **60**, 97–102.

Holding, D. H. Guidance in pursuit tracking. *J. exp. Psychol.*, 1959, **57**, 362–366.

Jones, R. E., Jr. Effects of delay of informative feedback, post-feedback interval and feedback presentation mode on paired associates learning. Unpublished doctoral dissertation. Univer. of Utah, 1963.

Jones, R. E., Jr., & Bourne, L. E., Jr. Delay of informative feedback in verbal learning. *Canad. J. Psychol.*, 1964, **18**, 266–280.

Kendler, H. H., & Kendler, T. S. Vertical and horizontal processes in problem solving. *Psychol. Rev.*, 1962, **69**, 1–16.

Keppel, G., & Underwood, B. J. Proactive inhibition in short-term retention of single items. *J. verb. Learn. verb. Behav.*, 1962, **1**, 153–161.

Lloyd, K. E., Reid, L. S., & Feallock, J.B. Short-term retention as a function of the average number of items presented. *J. exp. Psychol.*, 1960, **60**, 201–207.

Mandler, G. From association to structure. *Psychol. Rev.*, 1962, **69**, 415–427.

Miller, L. B., & Estes, B. W. Monetary reward and motivation in discrimination learning. *J. exp. Psychol.*, 1961, **61**, 501–504.

Murdock, B. B., Jr. Short-term retention of single paired associates. *J. exp. Psychol.*, 1963, **65**, 433–443.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. *J. exp. Psychol.*, 1959, **58**, 193–198.

Saltzman, I. J. Delay of reward and human verbal learning. *J. exp. Psychol.*, 1951, **41**, 437–439.

Spear, N. E., Ekstrand, B. R., & Underwood, B. J. Association by contiguity. *J. exp. Psychol.*, 1964, **67**, 151–161.

Spence, J. T., Lair, C. V., & Goodstein, L. D. Effects of different feedback conditions on verbal discrimination learning in schizophrenic and nonpsychiatric subjects. *J. verb. Learn. verb. Behav.*, 1963, **2**, 339–345.

Underwood, B. J. Ten years of massed practice on distributed practice. *Psychol. Rev.*, 1961, **68**, 229–247.

Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning.* Philadelphia, Pa.: Lippincott, 1960.

Wallace, J. Concept dominance, type of feedback and intensity of feedback as related to concept attainment. *J. educ. Psychol.*, 1964, **55**, 159–166.

Weinberg, D. R., Guy, D. E., & Tupper, R. W. Supplementary report: variation of post-feedback interval in simple motor learning. *J. exp. Psychol.*, 1964, **67**, 98–99.

Wilkinson, R. T. Interaction of lack of sleep with knowledge of results, repeated testing, and individual differences. *J. exp. Psychol.*, 1961, **62**, 263–271.

## Chapter 7

# Retention[1]

EDWARD A. BILODEAU
*Tulane University*

### HISTORICAL COMMENT

Until a few years ago my views on memory were based on Ebbinghaus (1885) and his successors, including Luh (1922). Improvements on the honorable and old tradition of Ebbinghaus seemed desirable. The apparent difficulty was that the shape of the curve of forgetting, long thought to be well known, was not really well described for human beings. Most modern research had dealt with acquisition curves, not retention curves. The field had been in doldrums for decades and neither independent nor dependent variables had been evolved or re-examined as in the field of acquisition. The prevailing opinion was that retention fell very fast in verbal learning and that retention hardly failed at all in motor learning. There were exceptions to the majority view; data obtained by McGeoch (1932) led him to contradict the generalization that skilled motor acts are better retained than are verbal materials and Hunter (1934) spoke of the failure to recall a motor skill minutes after a correct execution and went out of his way to state that there is no one curve of forgetting. The several studies of forgetting after overlearning explicitly or implicitly hold that little or no forgetting occurs in the immediate period thereafter.

---

[1] The principal part of this chapter is devoted to a discussion and overview of work originating in the Tulane Laboratory. The data cited result from the processes of memory observed in 11,000 subject sessions held during 1959–1964. The work of other laboratories is reviewed where it seems especially relevant, but does not get the prominence it deserves.

The present chapter will be more intelligible if Chapter 6 is read first.

Until recently, research in *verbal* learning has been fixed upon short-term memory and given over to the study of variables suggested 50–75 years ago. This situation changed rapidly with the work of Peterson and Peterson (1959), and Underwood and Postman (1960), who have introduced major new techniques and ideas to alter our thinking on portions of the curve for verbal materials. Indeed the whole field of memory has now witnessed an explosive growth, some of it being reviewed by Melton and others in a special issue of the *Journal of Verbal Learning and Verbal Behavior* (1963, pp. 1–119) and in other places (Bilodeau & Bilodeau, 1961; Naylor & Briggs, 1961; Bilodeau & Levy, 1964; Postman, 1961a, 1964). Important project work employing new tasks is being done by a number of productive investigators: Broadbent (1963), Bousfield (Bousfield & Puff, 1964), Cofer (1965), Deese (1961), Lloyd (Lloyd & Johnston, 1963), Murdock (1963), Reid (Reid, Brackett, & Johnson, 1963), Shepard (Shepard & Chang, 1963), Tulving (1964), and Yntema (1963).

Many recent developments can be traced to Underwood's 1957 review, where it was asked why the classical curve in verbal learning fell so fast and so far. After making a deduction for the contribution of proaction to the amount forgotten, he stated that verbal materials were much less subject to forgetting than earlier investigators had indicated. Underwood calculated that, instead of forgetting the greater part of a list of words within a few minutes, an individual might retain as much as 90 percent after 24 hours. This allowance for proaction, of course, altered our orientation toward the classical retention curve because most experiments on rote memory (with the exception of Tsai, 1927) had used the same subject over and over. Underwood's analysis found that repeated use of the same subject can be a methodological flaw unless one is interested in the phenomena of proaction for themselves. Since most verbal experiments of the earlier era have used this design, their data are limited to that use. Subjects who are given many lists to learn develop positive transfer in learning; in relearning they do poorly at first (negative transfer) and then regain original proficiency rapidly (positive transfer).

In the area of motor-skills learning, it is agreed that most studies have failed to find forgetting (for example, Lewis, McAllister, & Adams, 1951; Reynolds & Bilodeau, 1952; Battig, Nagel, Voss, & Brogden, 1957; Ammons, Farr, Bloch, Neumann, Dey, Marion, & Ammons, 1958; Mengelkoch, Adams, & Gainer, 1958; Fleishman & Parker, 1962; Roehrig, 1964). The failures made it impossible to determine the variables that produced more or less of it. Some has been found recently and now it can be said that there are portions of curves in motor learning available for examination and variation. The situation is not bright, however, for there is not much evidence in motor skills. In verbal learning, there are more data, but their meaning is in serious question.

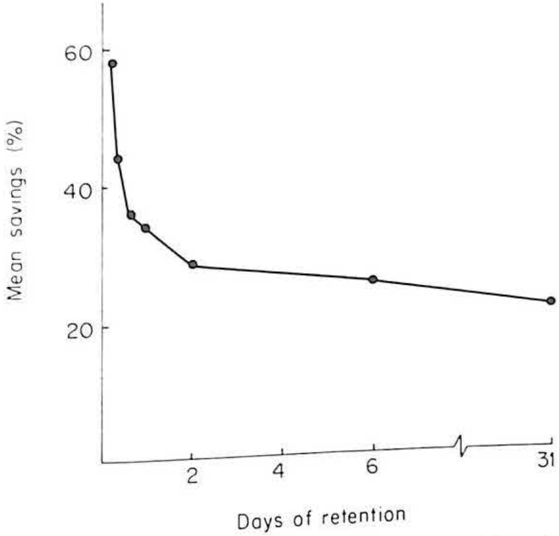


FIG. 1. An example of a common concave retention curve (data from Ebbinghaus, 1885: see 1913 translation).

### RETENTION CURVES FOR VERBAL TASKS

Most of us have been taught that the idealized curve is like the one shown in Fig. 1, which is plotted from data collected by Ebbinghaus. Here, a measure of savings in number of syllables recalled, based upon relearning, is plotted against the duration of the retention interval. The curve has a concave appearance and looks exponential. Accordingly, the classical saying is that forgetting proceeds rapidly at first, and later more slowly. As we have seen, however, Underwood's 1957 review suggested that a major portion of the bow was attributable to uncontrolled proacting agents. More recently, Bahrick (1964, 1965) has added other grave reservations and states that the shape is an artifact; it is the result of several factors, only one of which is the rate at which associations weaken in time.

Several methods of measuring retention, including recall, recognition, and anticipation, have been devised and compared. All have been criticized at one time or another, with Bahrick the most recent and articulate in questioning their meanings. Bahrick concludes that the entire area suffers from overgeneralizations and misunderstandings. In a series of persuasive arguments, he identifies three sources of trouble in the measures of retention.

#### 1. *Overlearning*

Especially in comparisons of recognition and recall memory, the recognition group ordinarily overlearns the material because both groups are given equal numbers of training trials. This may explain why recognition groups do so well in retention. Another common study uses a transfer design which may involve a penalty for a group changing from one method of performing to another. Here, two groups could be trained with $n$ trials by the method of anticipation, but tested for retention under recognition or recall. Bahrick warns of the limitations involved when the design calls for changing methods from training to test and training each subject to the same criterion of performance.

### 2. *Difficulty of Materials*

Though the foils in multiple choice tests in academic courses are usually weak competitors for the correct choice, this need not be so. Bahrick and Bahrick (1964) varied the level of difficulty of recognition alternatives and found, as they expected, that the easy recognition group was superior to the anticipation group, but the difficult recognition group was inferior. Recognition is not necessarily easier than other ways of retrieving events.

### 3. *Changing Sensitivity of Dichotomous Scores*

In the great majority of researches, the recall of a word is observed or it is not, the measure being dichotomous. According to Bahrick, estimates of associative strength as measured by anticipation, recall, or recognition will necessarily differ among themselves. Further, a measure changes in sensitivity as association strength changes in time. The measure is insensitive while every item is recalled, gains in sensitivity with decreases in associative strength until half the items pass below the threshold, and decreases in sensitivity as additional items pass below the threshold. Bahrick sums this up saying, "The slope of the curve does not directly reflect the amount by which associations are weakened in time. The slope of the curve simply reflects the number of associations passing below the threshold per unit of time" (Bahrick, 1964, p. 193). Comparing slopes of curves obtained under different methods or circumstances is hazardous because at retention time $(t)$ the methods may have quite different sensitivities. Even more dangerous to interpret are the relative amounts of forgetting within measures from one period of time to another.

The panels of Fig. 2 portray Bahrick's argument for three possible shapes of the distribution of associative strength and three positions with respect to the threshold of the indicant; the three curves in the lower portion represent the resulting forgetting-curve shapes. In another analysis he also showed how the variance of the distribution of associative strength in relation to the threshold of the indicant produces marked changes in the sensitivity of dichotomous scores. The distinction between responses observed and as-

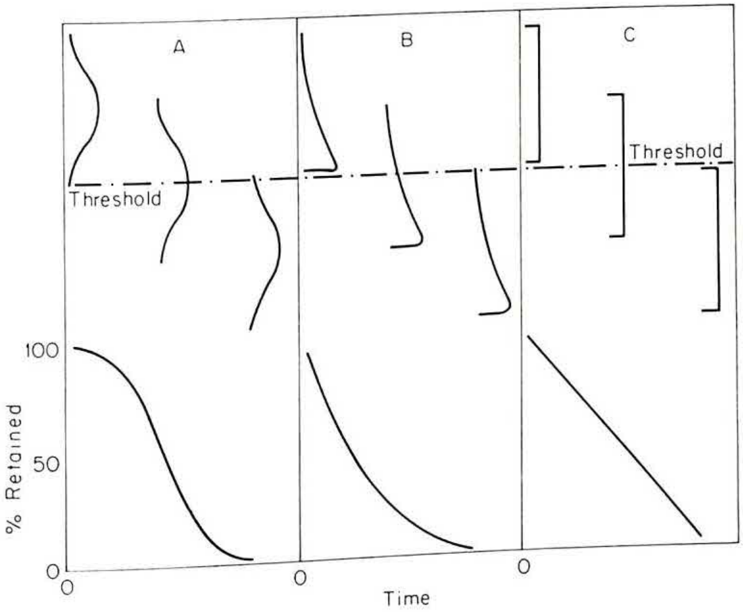


FIG. 2. Panels showing how three retention curves result from three different assumptions about the shape of the distribution of associative strength with respect to the threshold (after Bahrick, 1965).

sociation strengths assumed made explicit in Fig. 2 is a good one, for heretofore some psychologists talking of the forgetting curve have identified the curve of response with the strength of association itself, an error not only in mixing levels of discourse but also in setting one thing equal to another. The memory-performance distinction is analogous to the learning-performance distinction. In the case of retention of dichotomous measures we should be most reluctant to say that the events plotted are actually the associations themselves.

There are still other reasons why the literature fails to tell us much about the observed or the theoretical curves of forgetting: (1) Much of the older data cannot stand examination by contemporary research design, as a glance at Luh's (1922) section on Subjects will illustrate. (2) The duration of rest is not a variable in most forgetting studies. (3) Much research is directed to short-term processes and a shape is difficult to assess by extrapolation. (4) Many studies make use of two values of rest, too few to establish a shape. (5) The measurement of recall has evolved very slowly and curves for forgetting component (single) events within trials are wanting; their importance, however, is explicated in the following section.

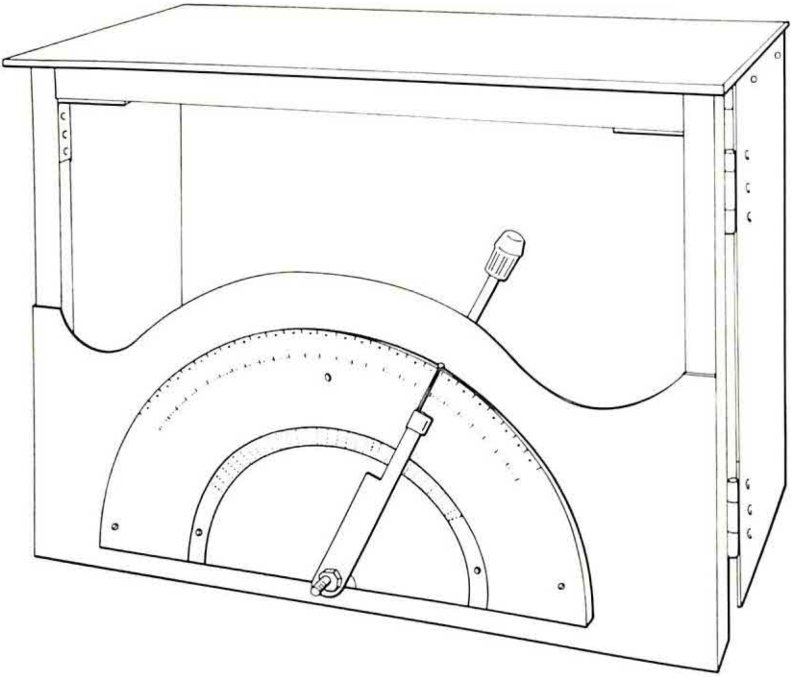


FIG. 3. An experimenter's view of the Manual Lever Apparatus showing the lever within the box and a scale for measuring amplitude of lever displacement outside the box (from Bilodeau, Jones, & Levy, 1964).

## RETENTION CURVES FOR MOTOR TASKS

Figure 3 shows a motor-skills device or lever task for learning a positioning response similar to line drawing. It served for a series of studies on lever-positioning, to be surveyed in detail below, and must be carefully described. The subject reaches into the box when told and moves the lever (blind) through several degrees. After the move is completed, the experimenter records the amplitude of response, and says something to the subject such as, "You were 54 units short of the target." The subject, then, has made one response called $R_1$ and the experimenter has administered one feedback or one knowledge of results which is designated KR. The lever shown in the figure was employed extensively for studies of feedback and its effect on learning (E. A. Bilodeau & Bilodeau, 1961). All kinds of numerical and temporal transformations of feedback were investigated in a program resembling schedules of reinforcement and with much the same objective – to see how changes in response depend upon the conditions of feedback.

In a series of studies on temporal delay, it was found that the point of interpolation of feedback between $R_1$ and $R_2$ had no effect on learning whatsoever (E. A. Bilodeau & Bilodeau, 1958). Feedback was delayed by seconds, minutes, hours, days, and weeks, and even so, learning was not dependent upon its temporal position. Instead it was found that the longer the time between $R_1$ and $R_2$, the poorer the learning. To make a long story short, the explanation of the poor learning with spaced responses held that subjects with long intervals between responses were forgetting some things of importance. Three basic events of training were being forgotten, it was thought. These events were $R_1$, KR, and also any planned response or response alternative ($R^a$) the subject might have had at the time of training. So it was that a project on learning and feedback turned out to produce some questions on memory. A *verbal* component, of course, was present from the start because motor studies typically use verbal feedback such as "long" and "short" and numbers such as "27" and "54." These feedbacks are often administered as *some* function of a previous response, and indeed, are often quite arbitrary and preprogrammed. This transformation is done in order to examine the relationships between KR and its response associates. In a simple case, for example, a KR of *32* (purporting to be the response error in a task such as the lever) should yield a lever displacement different from a KR of *324* and generally the former would produce a smaller movement change the next time. There ought to be both levering and numerical associative hierarchies to KR whatever its form or code – numbers, signs, units, etc. If the reader tends to make the error of equating the KR term with the response term, he should refer to Chapter 6 of this volume.

It was obvious that no one knew much about forgetting feedback, or for that matter, of forgetting the previous response alternative (E. A. Bilodeau & Bilodeau, 1961). These three component events of a trial had been ignored, confounded, or else left unelaborated in the past. In the standard experiment on retention, the subject was asked to do his best at recall, and even overtly asked to improve as in relearning; he was not asked to repeat anything such as a component of training. Requesting him to improve did not set the subject the proper task. Furthermore, little attention was paid to introducing special stimulus associates into the recall environment. In short, on balance most of the operations characterizing a standard retention study were in common with the learning study. The high degree of similarity in procedure may have been a function of the prevailing conception of learning and forgetting, that is, one is the inverse of the other.

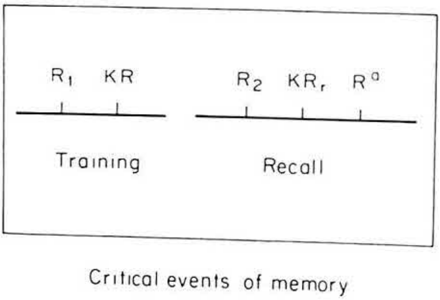


FIG. 4. Identification of the two major events of a trial of Training and the three major events of a trial of Recall (from Bilodeau & Levy, 1964).

To effect a remedy, the basic events which might be subject to forgetting in a standard learning situation were investigated (Bilodeau, Sulzer, & Levy, 1962). One answer is shown in Fig. 4 where positioning a lever will again serve for an illustration of the task. Here, $R_1$ is the first and only levering response of Training, $R^a$ is the alternative response or the response that the subject could make in place of $R_1$, and KR is the feedback pertaining to $R_1$ or to $R^a$. For example, if the subject chose to move to one place instead of the other of two places he had been previously shown, the move executed would be $R_1$ and the move not executed would be $R^a$. When asking for recollections of training events, the experimenter should probe for the recall events named $R_2$, $R^a$, and $KR_r$, respectively. The $R_2$ is, naturally, the subject's attempt to repeat $R_1$, and is to be performed regardless of KR or $KR_r$; $R_2$ is not an attempt to improve upon $R_1$ or an attempt to relearn $R_1$. The figure invites *inter*trial and *intra*trial comparisons and has proved to be extremely useful. Learning was looked for in increasing consistency in repeating responses, and forgetting in decreasing consistency in repeating responses.

The general objective in a series of four major studies involving the use of several thousand subjects was to relate the individual events of the recall trial to those of the training trial and to each other. There had been little direct investigation of these events before, or of the way to best express the relationships among them. Examining Fig. 4 leads one to ask a great number of questions. To take just one example: Will the subject's recall of $R^a$ depend more upon (a) the recalled feedback ($KR_r$) or (b) the administered feedback (KR). Incidentally, these questions help to show that the present analysis is not very task oriented and these remarks are directed as much to memory in general as to motor memory in particular.

The levering data show much forgetting of this simple skill, provided the data are examined in the right way. In the past, most analyses of memory could be reduced to the form: There is no difference between the average scores of Training and Recall. Yet we all know that means can remain constant while major changes occur in the consistency of scores from one distribution to another. Also, the mean may vary, yet still fail to describe the major changes occurring within a distribution of scores. Certainly forgetting is a far richer phenomenon than implied by a subject's averaging fewer responses in retention than in training.

### *Measures of Memory*

Changes between two arrays of scores such as produced by one trial of Training and by one trial of Recall can be measured by the mean, variance, and correlation. One, two, or three of these statistics can change in a total of seven combinations. Ordinarily, we find these seven combinations reduced to one and oftentimes the most trivial one at that. To wit, the average amount of material recalled is less than the average previously learned. If a subject correctly anticipates items 1–4 of a 10-item list during Training and items 5–10 during Recall, his retention score by simple tally exceeds the Training score even though there is no overlap on the items produced. Though this example is extreme, it brings out the paradox associated with counting the events of Training and Recall. The subject can be said to remember more than

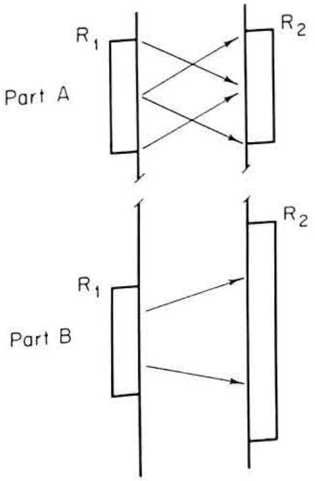


FIG. 5. In part A, $R_2$ represents a distribution of responses where the subjects attempt to repeat their prior responses ($R_1$). Here, forgetting has led to a correlation of zero, but mean and variance have held constant. In part B, forgetting has led to an increase in variance, but mean and correlation have held constant (from Bilodeau & Levy, 1964).

he learned (2 items more) or said to remember nothing (no items of Training repeated in Recall). If memory is to be viewed as persistence and forgetting as inconsistence, statistics of change and communality ought to be used.

Figure 5 shows how forgetting may occur without changes in the means. In part A, the distribution of the first response of a large number of subjects is shown under $R_1$. The distribution labeled $R_2$ represents the attempt of each subject separately to repeat $R_1$. The arrows show what might happen to four hypothetical subjects. In short, they do so badly at recalling $R_2$ that the correlation between the two arrays of scores ($R_1 \cdot R_2$) falls to zero. Some of this change may be attributable to experimental error and some to forgetting. By comparing the two distributions of the figure still further, we see clearly that the means and variances remain unchanged and are not sensitive to the forgetting taking place. Part B of Fig. 5 represents the case with a change of variance, no change in mean, and no change in correlation. Here, the subjects as a group are becoming more inconsistent than ever, yet $z$ scores and means remain the same from Training to Recall.

The ideas illustrated in Fig. 5 suggest that memory can be treated in terms of response consistency. It is for this reason that the more frequent use of variance, correlation, *or any other statistic* which will describe consistency and lack of consistency was urged (Bilodeau & Levy, 1964). This recommendation is easily applied to motor skills where the magnitude of the error of

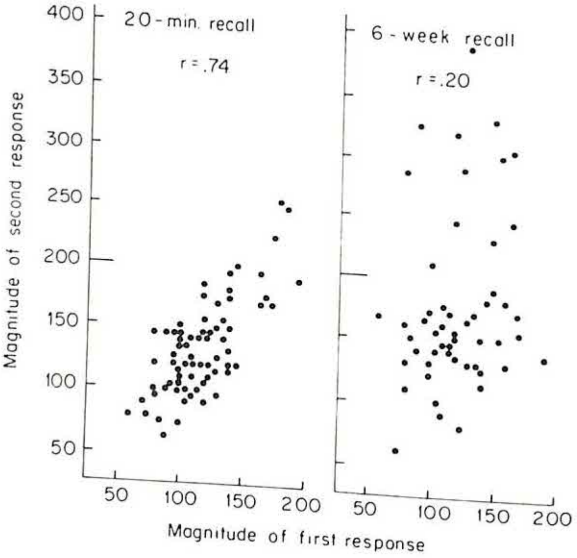


FIG. 6. Scatter plots showing the deterioration of $R_2$ with respect to $R_1$ after 20 minutes and after 6 weeks (after Bilodeau & Levy, 1964).

response is calculated as a matter of course (but I do not know what it means for traditional serial or paired-associates tasks).

*Data*

Figure 6 shows how forgetting $R_1$ is measured by plotting the magnitude of $R_2$ against the magnitude of $R_1$. The scatter plot on the left side of Fig. 6 shows forgetting after a 20-minute retention interval; the one on the right, showing greater deterioration, is for a 6-week period. In 6 weeks the correlation *(r)* dropped from .74 to .20.

Figure 7 is one from a number of studies of levering. The question is simply, "What happens to the variance of the second response over time?" In 6 weeks the variances rise by factors of 4 or 5, and, as usual for levering,

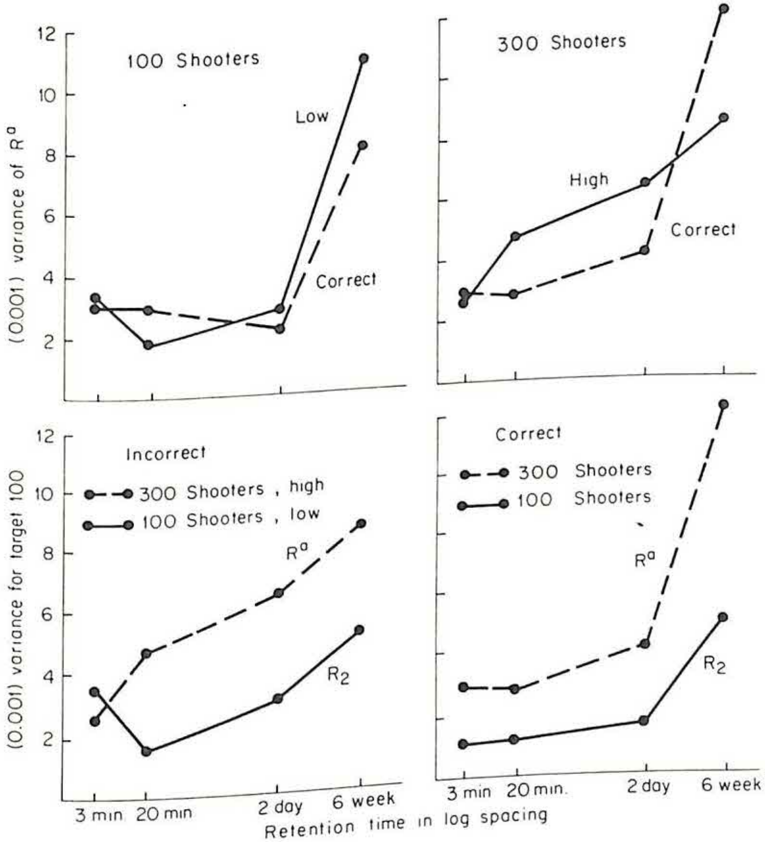


Fig. 7. Four illustrations of increasing variance in the recalled response as a function of retention time and other conditions (after Bilodeau & Levy, 1964).

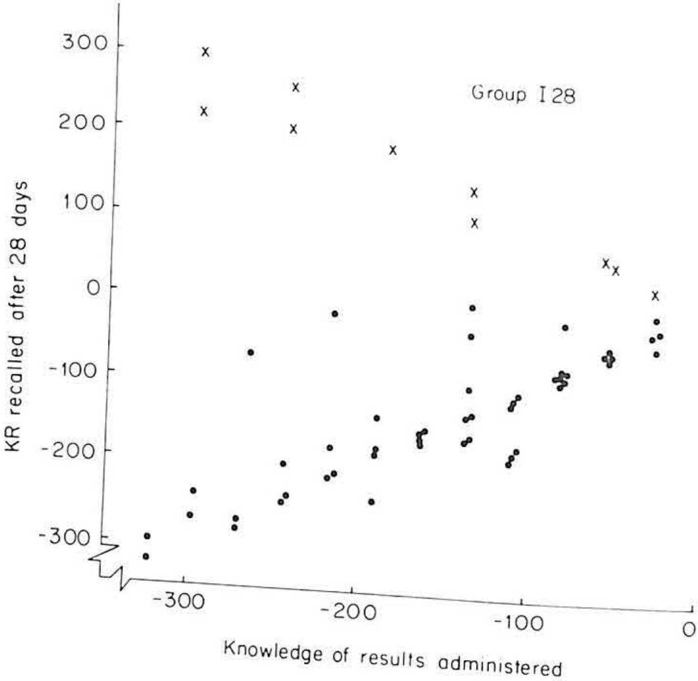


FIG. 8. Knowledge of results recalled ($KR_r$) has been plotted against knowledge of results administered (KR) for 60 subjects verbalizing a number and a sign. The X's represent the subjects who recalled the wrong sign (after Bilodeau *et al.*, 1962).

the variances take a positive acceleration; they change in the way shown especially well in the fourth panel.

Figure 8 shows how KR has been forgotten 28 days after its administration. The value of $r$ is about .30, showing how little there is in common between the number administered and the number remembered. Forgetting feedback has many aspects, including changing the absolute value, changing the sign, recalling the sign but not the number, and not responding at all. Moreover, the levering recalls are strongly related to $KR_r$, the recalled feedback.

Figure 9 is another sample from a number available. The correlations between $R_1$ and $R_2$ remain high and stable for at least 2 days of retention and then plummet.

Analyses of consistency of levering from Training to Recall have shown that individuals become more variable over time and that they change position relative to one another. Both of these changes occur in a lawful way. The means also change, but they did not appear to be the most important measure.

The over-all objective of describing the relationships between the events of Training and Recall and between the events of Recall themselves was

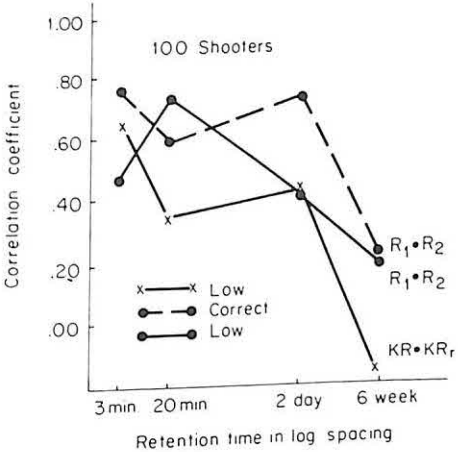


FIG. 9. Three illustrations of decreasing correlations between the events of Recall and Training with increasing value of retention time (from Bilodeau & Levy, 1964).

accomplished, first, under conditions of free recall as described above, then under conditions of modified free recall as will be described in the section on Controlled Recall (p. 333). That is, we soon began to stimulate the recall of one event before asking for the recall of another event. The first part of this chapter is devoted to free recall, the second to controlled recall. Now it will be necessary to examine more closely how a given association fares over time.

### *Analysis of Shape*

The curves shown in Figs. 7 and 9 are not like the ones we have been taught to expect from Ebbinghaus or from learning theory.[2] Perhaps the reason is simple; variances and correlations plotted on the ordinate are not common because we have not acted as if the components of memory were a facet of response reliability.

A proper way to set up a retention curve at $t_0$ is to establish the prior condition of some known degree of response consistency from one moment to the next. Having done this, we can begin to phrase our questions better: Given a known degree of reliability between one set of responses and a second set obtained a short while thereafter, what happens after that? We have seen from the lever series that the answer is: Nothing much happens for awhile, then the relationship deteriorates rapidly. In other words, forgetting proceeds faster and faster.

---

[2] The shape of the curves in Figs. 7 and 9 is unaffected by taking the logarithm of time. Logarithms were used so that both ends of the curve could be well seen by eye.

After examining a great deal of lever data, Bilodeau and Levy (1964) proposed a two-stage curve of forgetting as one of the curve shapes in retention. Looking back at the data plotted in Figs. 7 and 9 we can detect some forgetting taking place even before the first retention test. This early forgetting was identified as Stage I, and the later, negatively accelerating kind of forgetting, as Stage II. Stage II is most unlike the Ebbinghaus curve. In the illustration shown in Fig. 10, the intercept *(Y)* has been fixed at unity (rather arbitrarily) so that at $t_0$ recall is perfect. Recall is perfect in the sense that the ratio of the variances for $R_1/R_2$ is unity or that the correlation for $R_2 \cdot R_1$ is unity. The communality in variance $(r^2)$ between Training and Recall can be estimated from the figure, thereby increasing its value.

Certain important activities were usually interpolated between Training and Recall even when the interval was only 3 minutes long. All subjects released their grip on the lever, moved the resting arm away from the top of the lever box, changed foot position, moved out of the room, and moved back into position in front of the lever. These events occurred between $R_1$ and $R_2$ and under these circumstances the correlations have been about .70. In an earlier study some subjects did not disengage themselves from the training environment and these correlations were over .90. It seemed that the acts of moving away after $R_1$ quickly dropped the correlation about 25 units and nearly doubled the variance, an hypothesis later checked out by evaluating disengaging and nondisengaging treatments (Boswell & Bilodeau, 1964).

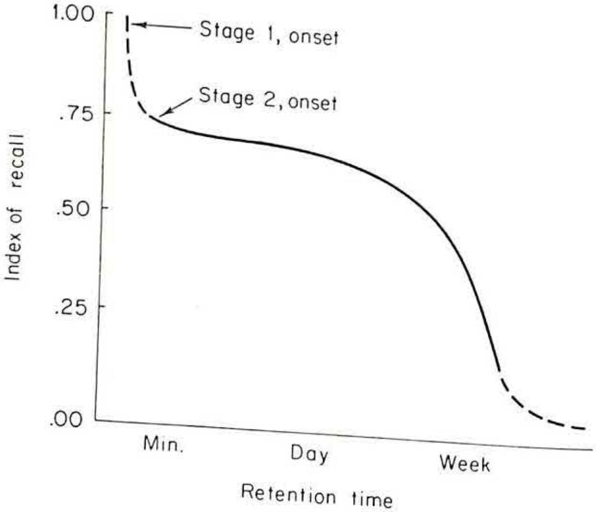


FIG. 10. Two stages of forgetting hypothesized by Bilodeau and Levy (1964). The longer, solid section is called Stage 2 and is preceded by a brief period called Stage 1.

Irion (1948) and later Adams (1961) made thorough analyses of the ways in which the loss of set could produce the first one of the two portions of retention loss. Disengaging the subject could produce (a) responses that interfere with the recall of $R_1$ or (b) changes in postural and environmental cues used by the subject to reproduce $R_1$ more exactly. If the changes in stimuli and responses associated with disengagement took rather sudden effect and then no more, the first stage could be explained as a warm-up phenomenon or as a loss of set.

Perhaps on occasion the retention curves for rote memory have also been dealing with Stage I through the action of disengagement or loss of set. Stage I might also be deepened by proaction, as supposedly happened to Ebbinghaus, who abused himself with the repeated use of training material. Data by Keppel and Underwood (1962) suggest the concave bow to be absent when proaction is a minimum and other data too have indicated relatively little forgetting in short-term verbal memory when naïve subjects are used.

Stage II is novel and its convexity is in great need of elaboration and rationalization. Data supporting the concave Ebbinghaus function are generally explained by referring to exponential-like decay processes. This explanation is popular and seems natural because so many different kinds of data can be fitted by the exponential. If the idea of reliability is used, however, the unique naturalness disappears, for there is an attractive way of rationalizing any convexity whether for Stage I or II. Given that the subject can repeat a response from one moment to the next, he should be able to repeat nearly as well if the attempt at repetition is delayed by an extra moment. Said another way, if the probability of repeating a response after an instant is high, there is no reason to assume that something unusual will happen during a second instant. The probability of repetition across the two instants should also be high though somewhat less than for the first one. Perhaps the starting of a car can be used for an analogy to the question of response reliability over intervals $t$ and $\Delta t$. Assume the car is started but once a day and that no failures at starting are ordinarily met with. What is the probability of failure with the interpolation of one extra day? It is greater for two days than for one to be sure; the greater increment in probability should occur in the second day of layover.

The above amounts to a situation without the deliberate introduction of proaction or loss of set by the experimenter, or a situation where, because of positive transfer, they have been overcome by the subject. Both retroaction and proaction can be worked into this probability model of the second stage even though there is no obvious precedent in memory theory to account for the delayed interference of previous events or of interferences yet to be encountered. In other sciences there are examples of effects temporally displaced from their initial cause. There is the odor trend of a rotting egg, a process of growth set in motion by earlier events. Only a trifle of updating of pro- and retroaction concepts is required and the updating is based on the general idea that the probability of the occurrence of interfering events should rise in keeping with the expectation of events randomly distributed in time and that the growth of their effect is ogivelike. How this might be set up for both proaction and retroaction is described below.

(1) Proaction: Past experiences can act upon present response reliability, not necessarily immediately and totally as suggested by a concave curve, but corrosively as suggested by convexity in the curve. The reason might be that the probability of proactions or their effects rises in time in a fashion analogous to the ogivelike accumulation of chance events.

(2) Retroaction: The effect of new experiences upon response consistencies established in the past is not necessarily total and immediate, but can be rather slow and accelerating. The reason might be that the probability of encountering retroactions and their effects rises in time in a fashion analogous to the ogivelike accumulation of chance events.

There are no compelling data to suggest that suggestions 1 and 2 above on pro- and retroactions are correct. We can anticipate several different models which could account for the various shapes of retention curves now being discovered and it is urged that we begin to think about how one might account for shapes different from the classical one. The following pages take a much closer look at the model of random events and their effects.

### *Repeated Recalling*

In the studies considered so far the subject retrieved an event just once; there was no repeated recalling. As much as 6 weeks intervened between $R_1$ and $R_2$, but there was no further call for the repetition of $R_1$. If $R_2$ had been scheduled earlier (say Day 2) and a third response $R_3$ at 6 Weeks, we might wonder about the effect of the earlier retrieval on the later retrieval, i.e., $R_3 = f(R_2)$. An elderly subject reminiscing about earlier events may actually substitute more recent recalls for the earlier original events. Out in the street or in the laboratory, then, a recollection might be traced to a previous recollection, to the original event, or even to both. Bilodeau, Jones, and Levy (1964) initiated the examination of this substantial question. The matrix of intertrial correlations for learning tasks was used as a point of departure.

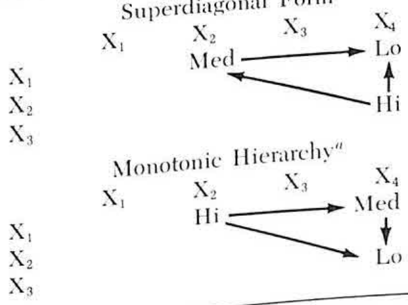
Reynolds (1952), Fleishman (1953), Jones (1962) and others have cast up a great many matrices for motor-skills learning tasks. These tasks usually involved learning to stay on target as with rotary pursuit apparatus. Jones summarized and evaluated this literature, concluding that a particular pattern of *r*s, the *superdiagonal form*, is universally observed. In this pattern a trial correlates most strongly with its neighbors, most weakly with the first and last trials of the series, and so on, as shown schematically in the upper part of Table I. The superdiagonal pattern is obtained under orthodox training procedures. That is, the dominant theme for the subject is to obtain a better and better score from trial to trial. Especially at first, if the subject were to repeat an earlier response, it would imply a partial failure of learning. Rather than repeating, the task is to change the response further in order to reduce errors, to maximize time on target, etc. Toward the end of training, as learning approaches its limit, the process of response change may be gradually arrested and the process of repeating responses may be replacing it. Here, then, in the latter part of a training session *repetition* instead of *change* characterizes the process. The situation is actually a complicated one, for many training periods should be expected to show both acquisition-like behavior and recall-like behavior (see also Murdock, 1960; Tulving, 1964).

Repetition, it should now be recalled, is the key word for the lever series of retention studies. It is important to distinguish between the procedure of repeated recalling (retrieving or retention testing) and of repeated practice (storing or acquisition training), the repeated recalling referring to repetitions with no improvements required and the repeated practice to a series of improvements with no repetitions required. In an experiment on the effects of repetition *qua* repetition (Bilodeau *et al.*, 1964), the subject was instructed to repeat the training response (now to be called $R_0$) not once, but four times. The retention intervals for successive repetitions (called $R_1$, $R_2$, $R_3$, and $R_4$) were 3 minutes, 20 minutes, 2 days, and 6 weeks, respectively.

The resulting pattern of *r*s is shown in the lower half of Table I, and it was a curious one, the *monotonic hierarchy*. Neighboring trials do *not* correlate

TABLE I

Matrix Patterns for Acquisition and Retention Trials



[a] Note, for example, $X_1 \cdot X_4 > X_3 \cdot X_4$

highly, except for the earliest repetitions, and the last two terms of the series produce the lowest $r$. According to Jones (1962), the monotonic hierarchy cannot be rationalized in the same way as the superdiagonal pattern. The latter he interprets as the product of a process of simplification. That is, progressively fewer factors are required to explain the variances over a series of learning trials. But for the monotonic hierarchy, only one factor is required at the start and less and less of it is used to explain later attempted repetitions. The data showed *the* one factor to be the recall of $R_1$ (note, not $R_0$) and its progressive dilution was attributed to increasing amounts of accumulating random variance as discussed earlier. Thus, it was concluded that the repetition process was neither the same as nor the inverse of the learning process. The repetition process did not appear to be one of unlearning. A study by Barnes and Underwood (1959) was also directed to a similar problem.

The conclusion, it should be noticed, is not directly tested. It is based on lever data for a repetition schedule and rotary pursuit data (and the like) for an acquisition schedule. It is not certain that repetition was the sole variable responsible for the monotonic structure. The conclusion is more a conjecture than a well-established principle. The subject of patterns of $r$ seems especially promising and the inquiry should start by looking for the training, interpolated task, and retrieving variables that change the patterns. One very recent experiment (Bilodeau & Blick, 1965) intercorrelated retention trials for verbal materials, and here too, the pattern was monotonic. Since sources of interference were deliberately inserted into the retention environment, it was possible to conclude that progressive complications characterized the retention process. This experiment is discussed in greater detail on pages 342–345.

One other finding was of general importance, this pertaining to the matrix for successive recalls of KR. The KR matrix tended toward the superdiagonal form; neighboring trials seemed more alike than more separated ones, and the highest $r$s were between retrievals 3 and 4. It could be that forgetting verbal materials like numbers (a subject might recall 324 as 324, 320, 300, 300 over four successive trials) is more systematic than forgetting levering, which is dominated by accumulating random variance. However, not much is known and so not much can be explained now. Again we see the need for an extensive inquiry where the manipulation of patterns in the matrices for learning, relearning, and repetition is the objective. For additional discussion of intertrial structures see Jones, pp. 112–127.

## SUMMARY

Many dependent and independent variables of memory have been neglected and this has hampered the discovery of curve shapes different from the classical one. Until recently, there has been no curve to talk about for response repetition, response alternative, or knowledge of results. The importance of variance, correlation, and other statistics of change should receive more attention as ways of measuring forgetting. Thus I have stressed memory as a question of persistence or response repetition, and forgetting as a question of response change. Response change implies more than is conveyed by a tally of events retained. At least some forgetting seems a process different from learning and attributable to accumulating variance acting through proactive and retroactive sources.

## Controlled Recall

### Historical Comment

Research on retention in the Tulane Laboratory actually began with lever positioning experiments and the effects of delay of verbal feedback. This activity shaded off into studies of the consequences of infrequent responding and infrequent feedback, effects that belong to the domain of forgetting. An impetus to a series of studies of verbal memory came from a desire to make a demonstration that the lever procedures had verbal analogs. Indeed, the verbal analogs turned out to be so workable that the motor series has been suspended for a while.

The first strictly verbal investigation was performed in 1958 and is called the *sweet-ugly* study because the two words *sweet* and *ugly* were shown side by side for a few seconds on a screen. The subjects were told: "One of these two words is correct, write your selection on your card." The word written by the subject was defined as $R_1$; the word witnessed but not written down was defined as the alternative or $R^a$. Feedback was administered when the experimenter said, "You chose the correct (or wrong) word." The subject was not instructed to learn or remember *sweet, ugly*, or correct (or wrong). Five weeks later the experimenter appeared again and asked for the recall of *sweet, ugly*, and correct or wrong and found that all were forgotten to some extent and that the recall of one was dependent on the recall of the others.

The *sweet-ugly* experiment made good use of stimulated recall. Two of the three events occurring during Training were readministered by the experimenter during Recall and only then was the subject asked for the third event. For example, "Five weeks ago, you chose the word *sweet* and you were correct. What was the other word?" It was in this way that memory for one event was measured after a reinstatement of the memory for the others. An elaborated stimulation methodology is reproduced in Table II. It shows the eight environments where forgetting may be found and these range from unstimulated recall (free recall) to completely stimulated recall. Rows 1, 2, 3

TABLE II
A CLASSIFICATION OF THE POTENTIAL TYPES OF FORGETTING SITUATIONS[a]

| Event reminded | Potential forgetting |
|---|---|
| $R^a$, KR | $R_1$ |
| $R_1$, KR | $R^a$ |
| $R_1$, $R^a$ | KR |
| $R^a$ | $R_1$, KR |
| $R_1$ | $R^a$, KR |
| KR | $R_1$, $R^a$ |
| None | $R_1$, $R^a$, KR |
| $R_1$, $R^a$, KR | None |

[a] After Bilodeau *et al.* (1962).

show two events used as stimulators, Rows 4, 5, 6 show one-event stimulation, Row 7 shows none, and Row 8 shows three. The *sweet-ugly* study used Rows 1, 2, 3 of the table for the experimental variable. An earlier experiment (E. A. Bilodeau & Bilodeau, 1958) in a motor study used one reminder (KR, Row 6). The experimenter said, "When you were here yesterday, I told you that your move was low by *X* units. Now, let's try to improve upon yesterday's response." This instruction to the subject is a reminder of KR and a probe for the alternative response. The KR-reminder design now takes on the appearance of a small part of a much larger problem which is the theme of this section.

*It was thought that the forgetting of one event would bear on the forgetting of others. Conversely, the recall of one event (even artificially aided recall) would bear on the recall of the others.* The amount of bearing, it was hypothesized, should depend upon the strength of association between events. That is, the associative structure of the stimulus would be the critical knowledge on which to base the prediction.

The events $R_1$, $R^a$, and KR probably become weakly interassociated during training for they are in close temporal contiguity. As in learning a list of paired associates, there are probably two types of associative processes occurring simultaneously; (a) the learning of the events *qua* events (responses or items) and (b) the association between events. With rest there is a weakening of both types of learning. If the subject is reminded of one event (or two), the remaining event (or events) can be more readily produced, as the experimenter takes advantage of (or cues) the associative connections among events, hence the presumption that the recall of a given event is partly contingent on the recall of others. A reminder of $R^a$ and of KR would serve to arouse $R_2$ because $R^a \rightarrow R_2$ and KR $\rightarrow R_2$. But an $R_2$ is not ordinarily elicited in this way. In the typical free-recall experiment, events $R^a$ and KR do not intervene or precede so explicitly and the ordinary antecedent to $R_2$ is merely the experimenter's question "What is $R_2$?" and not a string of explicit reminders. Outside of the laboratory, people retrieve information in the presence of a great deal of contextual background. In the laboratory this background is more often held constant than varied, and as a result, information about it is scarce.[3]

Another way of looking at the controlled use of stimulation during retention is in terms of experimental design. This has been stated as: ". . . we wanted to know what happened when we permitted the forgetting of one event and at the same moment the other two were recalled perfectly by each and every subject. Obviously, assuring the recall of the other two events is a form of holding relevant stimulus variables constant and should be a prized form of design. Most other experiments permit the subject to forget all, a design corresponding to the seventh row of the table" (Bilodeau, Sulzer, & Levy, 1962). The *sweet-ugly* study, which reminded of two events, was done in the lever context and was patterned after the *sweet-ugly* procedure (Bilodeau, Levy, & Sulzer, 1963a). The response was measured by its amplitude, and feedback was also expressed as a number or variable quantity. Thus, the training event is recalled in some measurable degree, variance and correlation measures express the degree of change across rest, and the difficulties involved with dichotomies (Bahrick, 1964) are avoided. It turned out that $R_1$, $R^a$, and KR were only weakly associated and this proved to be a shortcoming. The study is not especially valuable for its data, but contains an extended evaluation of stimulation methodology, both motor and verbal. In subsequent verbal studies our S-R terms were more strongly related and the effects correspondingly more gratifying.

---

[3] There remains the necessity of measuring the retention of $R^a$ and KR in more strictly verbal contexts than yet accomplished. The analysis of feedback, in particular, has been overlooked because feedback is so often confounded with the response term. In lists of paired associates, for example, *x-y* is exposed after the stimulus term *x*, *y* being ordinarily considered the response term. Indeed, it was observed long ago that *y* serves as a confirmation as well, but in fact, its informative aspects had been neglected. Bourne's comments on: Information Feedback (and especially his review of his work on *feedback modes*) in the preceding chapter, can be read most profitably, however, to show that progress is now being made in this neglected field.

In still another type of experiment, the $R_1$ term represents the response from one of several alternatives as in a multiple choice test. If items X, Y, Z were presented simultaneously, and the subject chose Z, it would be desirable to observe the forgetting of the alternatives (X and Y), and of the feedback (correct and incorrect) and their interdependencies.

## RECENT STUDIES OF VERBAL STIMULATION

Since I wish to discuss recent work on verbal retention, I return to *sweet* and *ugly*. They were selected after much vacillation and finally chosen from many word pairs because they are not associated; that is, *sweet* does not ordinarily evoke *ugly* and *ugly* does not evoke *sweet*. The choice of words with low interitem associative strength was probably a mistake. On the other hand, *black* and *white* might have been selected, but they were rejected because of high associative strength, and a consequent overly high susceptibility to reminding.

If *black* and *white* had been used as training items, many subjects with no recollection of *black* plus *white* might be found a year later. But were the experimenter to begin the retention session by saying, "Last year you chose 'black,' what was the other word?", the response "white" would surely be enhanced and memory facilitated. *White* is the dominant or primary free-associative response to the stimulus *black* and has been evoked by stimulation. The task has been changed from the difficult one of the unaided recall of *white* to the easier one of recognizing *white* because the administration of the stimulus evoked the associates of *black*. Because *white* is the primary associate, it stands a good chance of being recognized as an appropriate response. *By ascertaining the strength of the associations between a stimulus and its responses we would be in a position to predict occurrences of responses with a knowledge of the stimulus.* This is precisely the turn of events that ended the search for words to replace *sweet* and *ugly* (and lever studies too). The words we wanted were the stimulus and response items of the norms on free association collected by Russell and Jenkins (1954). However, Russell and Jenkins had not processed the strength of association between their stimulus and response terms sufficiently. Before going into these calculations in detail, let us examine how our efforts at regulating the events of recall relate to some other movements in retention.

In an extensive review of 50 years of work on rote learning, Underwood (1957) said that he would like to see retention studies on the forgetting of a single association. His wish came true, for Peterson and Peterson (1959) set off a train of studies involving the retention of one nonsense syllable *(cvc)* after one trial of training, and the *sweet-ugly* study initiated a similar series of motor studies also involving the limiting case for events and trials.

Shades of stimulated recall were used by Jenkins and Russell (1952) and Bousfield (1953) some time ago. The former broke up pairs of words consisting of a stimulus and its primary taken from free-association norms and trained the subject with both items in randomized order. For example, *black* and *white* would appear in the same list, but never as adjacent items. Recall was scored for the *re-pairing* of stimulus and response items imposed by the subject, or what is called "associative clustering." It is as if the recall of one item served as the stimulus for retrieving an associate, and so, perhaps, it could be said that the reorganization at Recall is attributable to the subject's providing his own stimulus or reminder (see also Deese, 1959a). Rothkopf and Coke (1961) used a similar explanation for the order imposed by the subject in the free recall of a large number of words of varying associative interdependencies. Free recall has always been operationally defined, and that is a good thing, for it has become quite clear that a great deal of what a subject will recall depends on what he has just recalled or said (Conrad & Hille, 1958; Murdock, 1963). Cofer (1965) provides a very recent review of variables dealing with training materials and how they influence the sequential organization of the materials recalled during free recall.

In stimulated recall, the stimulation can be administered in any number of ways. For example, S-$R_1$ pairs could be split, with the responses used as training items and the stimuli introduced as memory aids during the retrieving period (Bilodeau *et al.*, 1963b). Again, note that the stimulus is employed during Recall, not during Training. If the right sort of stimulation were used, the split-pair procedure could produce enhanced recall. To promote errors, stimuli which evoke associates to substitute for the training words, as can be done in recognition tests, could be introduced. These two ideas quickly spread into a search for stimulus rules to produce facilitation and for other stimulus rules to produce interference. The search began with an analysis of the free-association process itself. It was thought that rules of free association would be effective in foretelling what a subject would retrieve in a stimulated recall environment. After all, the method of free association *is* an elementary method of memory.

## Analysis of Russell–Jenkins Free-Association Hierarchies

For many years experimenters chose to work with nonsense syllables because *cvc*'s were thought to be almost without associates (nearly meaningless). During the 1950's, after Noble's (1952) pioneering work on meaningfulness, considerable scaling of nonwords and words took place. These quantities were used to predict one thing or another about verbal learning. Here too came the Underwood and Postman (1960) frequency hypothesis which holds that new connections are interfered with as older associates spontaneously recover with rest and take their place, words in more frequent use producing the most interference. This point of view finds ready expression in paired-associate learning methods where the stimulus and response terms may be either high or low in frequency of use, but where no special attention is paid to the pre-experimental strength of association between the S and R terms. Cofer (1958) and Deese (1961) have emphasized the importance of working with the pre-experimental associative strength between words. Garskof and Houston (1964) state that both association strength *and* associative overlap correlate with recall. Associative overlap refers to number of associates in common and the frequency of associates in common. The strength of association between S and R is at issue in the discussions of the levering data and of the *sweet-ugly* study. Research progress was swift and highly successful after suitable quantifications of associative structures of stimulus terms (Bilodeau, Fox, & Blick, 1963b; Fox, Blick, & Bilodeau, 1964). This amounted to obtaining pre-experimental information about word arousal and using the information to anticipate what the subject would retrieve in a recall session.

The Russell–Jenkins (1954) data show all responses given in free association by 1008 subjects to 100 common words. One of their word hierarchies is schematized at the left side of Table III. The cultural primary (now to be called $R_1$) is the single most frequent response to the stimulus word, the cultural secondary ($R_2$, not to be confused with a repetition of $R_1$ as used in the previous section) is the second most frequent response, and $R_3$ to $R_n$ symbolizes all remaining words in the hierarchy. In the illustration on the right, *chair* is the primary and its probability *(p)* of occurrence was computed to be .84. The *p* of the secondary *(food)* is .04, and all the words from *desk* to *company*, taken together, have a *p* of .12.

Presenting a word such as *table* for a stimulus during Recall instead of during free association would arouse all responses in the hierarchy weighted according to $p(R_1)$. This would be a good way of facilitating the retrieval of a primary, secondary, or other term in the hierarchy, if one of the terms had been exposed just before. If *chair* or *food* has been used as a training item, its recall is enhanced with *table* in the recall environment. In the first experiment, only secondary words were used for training material. Therefore,

TABLE III
A HIERARCHY OF DISCRETE FREE-ASSOCIATION RESPONSES ADAPTED FROM NORMS OF RUSSELL AND JENKINS

| General | Illustration | | |
|---|---|---|---|
| $S_1 \longrightarrow R_1$ | *Table* ⟶ | *Chair* | (.84) |
| $R_2$ | | *Food* | (.04) |
| $R_3$ | | *Desk* | (.02) |
| . . . | | . . . | . . . |
| $R_n$ | | *Company* | (.001) |

the stimulated subject was expected to say "food," when correct, and "chair," when incorrect.[4]

Since it had been decided to use the secondary terms of the hierarchies for

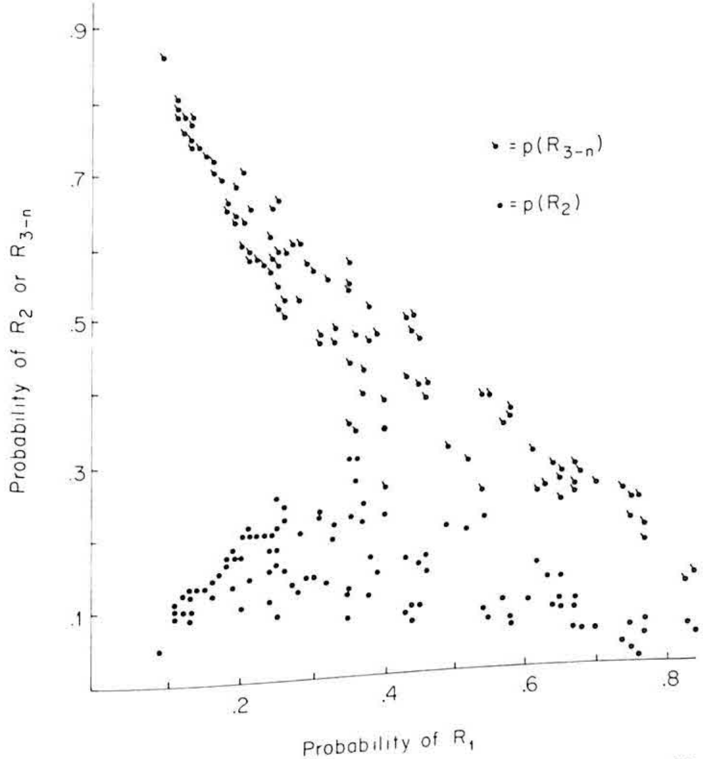


FIG. 11. Scatter plot showing the relationship between $p(R_2)$ or $p(R_{3-n})$ and the independent variable, $p(R_1)$, in the Russell-Jenkins hierarchies of associative responses. Note that $p(R_2)$ rises in magnitude and reaches a maximum when $p(R_1)$ is about .35 (after Bilodeau *et al.*, 1963b).

---

[4] The simplest experiment would expose *table* and *food* during Training, and later on, the subject would be asked to recall *food* in the presence of the stimulus *table*. Then we could compare the new and the old hierarchies. Since the subject has been trained with the secondary word (*food*), the primary word (*chair*) will be an especially strong competitor during recall. This study, however, was not done because it was decided to leave the stimulus term out of training. Instead, only response terms were to be used for training material; the stimulus related to them by way of the norms would be introduced during recall. We ended up using two training lists of five secondary words per subject and stimulated the recall for one list. The application of the stimulus was supposed to augment the probabilities of both primaries and secondaries over those of a group not so stimulated.

training material, it was thought necessary to examine the relationship of the probability of the secondary to the probability of the primary. A strong relationship between the primaries and secondaries was quickly discovered. This relation (eta equals .65) can be seen after careful examination of the 100 dots in the lower part of Fig. 11, a most important law being found. When $p(R_1)$ is very low or very high, $p(R_2)$ is low, as it must be; when $p(R_1)$ is of moderate strength, $p(R_2)$ reaches its maximum value. The best chance of responding with $R_2$ is with $p(R_1)$ in the neighborhood of .35. A strong and necessarily negative relationship between $p(R_1)$ and the probability of responses 3 to $n$ is shown by the 100 dots with tails in Fig. 11. It follows that three kinds of words or categories of words stand out in the figure.

*Category A:* Low $p(R_1)$, low $p(R_2)$, and high $p(R_{3-n})$.
*Category B:* Moderate $p(R_1)$, maximum $p(R_2)$, and moderate $p(R_{3-n})$.
*Category C:* High $p(R_1)$, low $p(R_2)$, and low $p(R_{3-n})$.

When used for stimulated recall, Category A words should maximize the incidence of responses 3 to $n$, Category B should produce the most $R_2$s, and Category C should produce the most $R_1$s. The mean associative $p$s for the categories are given in Table IV. The categories are names that represent the complex relationships or pattern among $p(R_1)$, $p(R_2)$, and $p(R_{3-n})$. Their use implies the circumstances for retrieval of correct words and also the circumstances for particular kinds of misrecall.

## CURRENT RESEARCH ON STIMULATED VERBAL RECALL

In the experiment that followed, the subjects were trained with two lists of five secondary words made up in a booklet. The lists were exposed for 20 seconds each and the instructions read, "study the words below." After studying the two lists, the subject canceled vowels for a 2-minute retention interval, then entered the recall period. During Recall he was asked to write

TABLE IV
RUSSELL–JENKINS NORMATIVE PROBABILITIES OF RESPONSE BY LIST CATEGORY AND POSITION OF RESPONSE IN HIERARCHY[a]

| Category | Position in hierarchy | | |
|---|---|---|---|
| | $p(R_1)$ | $p(R_2)$ | $p(R_{3-n})$ |
| A | .14 | .12 | .74 |
| B | .32 | .24 | .44 |
| C | .73 | .05 | .22 |

[a] After Bilodeau *et al.* (1963b).

down the words of one of the previous lists. It was found that recall with stimulation, as opposed to without, evoked more responses and led to large differences in the source and character of errors. The effect of category was strong, for the stimulation augmented the incidence of secondaries from Category B, augmented primaries from Category C, and augmented responses 3 to *n* from Category A, just as expected. The stimulation also led to corresponding decreases in intrusions from the Other List, Blanks, and Importations. It was concluded (a) that the stimulation evoked otherwise dormant responses (correct and incorrect ones) and (b) that changes in the category of stimulation varied the type of correct and incorrect recall according to word-arousal patterns in the language. That is, the free-association probabilities of the Minnesota norms classified by category accounted for the trends in an experiment involving training and recalling in other people 10 years later.

Using various categories of stimulus words is a way of expressing extra-experimental information about response arousal and setting up certain expectations about their proactive effects. Relating the primary, secondary, and all other responses to a common stimulus in the norms made it possible to employ that stimulus independently of the material of Training. A good many intrusions of an extraexperimental variety were produced, a phenomenon that Underwood and Postman (1960) and many others have also tried to demonstrate.

Few or about 4 percent of the many intrusions obtained were from the other list. This is about the level of confusion in list membership obtained by using any one of many list-building rules. Psychologists, it must be admitted, have not been able to produce much interference between lists (Underwood & Postman, 1960; Postman, 1961b). The 4 percent obtained seemed a good place to begin research on the deliberate stimulation of other-list intrusions. Blick (1965) undertook the task of making up lists and stimulation that would maximize intrusions, his objective being to increase them by an order of magnitude. He placed five primary words in List 1, five corresponding secondaries in List 2, and five corresponding stimulus words on the Recall page. As in the preceding study, the subject was asked to write down all the secondary words (List 2) he remembered. The logic behind the design was that experimental and pre-experimental associative tendencies for $R_1$ combine strongly and $R_1$ will be more likely retrieved because of the converging sources of interference. Alternatively, it could be argued that if the associative tendency for $R_1$ is very strong, it might be readily differentiated from $R_2$ and, hence, the subject will be reluctant to claim it for an $R_2$.

The results showed that many primaries were produced instead of the secondaries and the rate of intrusion went hand in hand with the tabled value of $p(R_1)$. The intrusions of obtained primaries for Categories A, B, C, were 22, 42, and 56 percent, respectively, showing that intrusions rise with increasing strength of association between stimulus and primary. Category C words produced 56 percent intrusions, meaning that most of the words recalled were from the wrong list—a finding that should be compared with the more usual less than 10 percent rate. Table V shows the intrusion probabilities by subject. The number 40 at the lower right side means that 40 percent of the subjects suffered four or five intrusions apiece (out of five). As a matter of fact, 25 percent of the subjects reproduced the wrong list of words in its entirety, and had no correct words, no blanks, no importations nor other types of misrecall. This shows that we know how to produce high levels of misrecall of known varieties. In effect, Blick has made good use of people's pre-experimental language habits and has led them to misremember in ways we can now understand. Cuing recall can be varied in several other ways to study the retrieval of stored information, but the promotion of intrusions from the other list was a good place to start. The general objectives were similar to Deese's (1959a, b, 1961), though he concentrated his attention on interitem relationships of the *training* material and tested for retention by the method of *free recall* (see also Coleman, 1963).

Up to this point it has been shown that the intrusion rate for primaries is abetted by exposing both primaries and secondaries a few minutes prior to recalling. An experiment with retention interval varied up to 28 days was undertaken to determine if increasing rest, but with no exposure to primaries, would not also increase the probability of primaries (Bilodeau & Blick, 1965). It was reasoned that training with $R_2$s raises their probability of occurrence, which then decays with increasing retention interval. Primaries (and $R_{3-n}$s) should substitute for secondary words as the subjects revert toward their normal free-associating state.

Figure 12 shows what happened to recall of $R_2$ over 28 days of retention.

TABLE V
PROPORTION OF SUBJECTS WITH INTRUSIONS OF PRIMARY WORDS[a]

| Category | No. of $R_1$'s 0–1 | 2–3 | 4–5 |
|---|---|---|---|
| A | .72 | .28 | .00 |
| B | .42 | .35 | .23 |
| C | .27 | .33 | .40 |

[a] After Blick (1965).

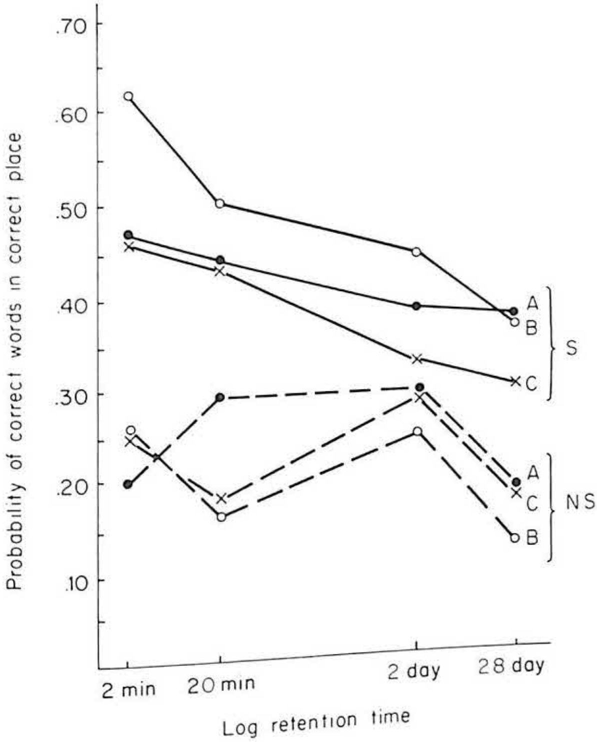


FIG. 12. Retention curves for correct words over 28 days with (S) and without (NS) recall stimulation, and three types (A, B, C) of stimulation (from Bilodeau & Blick, 1965).

The top three curves represent stimulated recall (S) using Categories A, B, C. The bottom curves (NS) show what happened to recall of $R_2$ when no stimulation was used. Obviously, the subjects did much better under stimulated conditions than under the not stimulated conditions and they also forgot more with retention periods beyond 2 minutes. Just as expected, the decrement in correct words retrieved was made up of an increment in $R_1$ s for Category C and $R_{3-n}$s for Category A as shown in Table VI. There was no increase in the number of unexplained words or of failures to respond even after 28 days. After 28 days the subjects were answering the stimulus more like free-associating subjects than before, but still, they fell far short of producing normative rates. The shapes of the curves are too ambiguous to allow for ready extrapolation, but they do permit the conclusion that the training

TABLE VI
OBTAINED RESPONSE PROBABILITIES FOR STIMULATED GROUPS

| Response | Category | Retention interval | | | |
|---|---|---|---|---|---|
| | | 2 Minutes | 20 Minutes | 2 Days | 28 Days |
| $p(R_2)$ | B | .62 | .50 | .44 | .36 |
| $p(R_1)$ | C | .25 | .30 | .29 | .40 |
| $p(R_{3-n})$ | A | .28 | .33 | .37 | .39 |

experience with secondary words had modified the subject's associative structure for a long time to come. This experiment is also valuable for showing that retention functions rise, remain constant, or fall depending upon the type of response measure, use of stimulation or no stimulation, category of stimulation, and length of the retention interval.

Intercorrelations of item analyses based on the probability *(p)* for correct response for each of the 45 words used in the above experiment, first under not stimulated conditions, then under stimulated conditions, are shown in Table VII and represent the first deliberate attempt to vary patterns of correlation from one recall condition to another. The matrix for S is different from NS in that (a) all the *r*s are lower and (b) there is a trend to the *r*s. The pattern for matrix S is monotonic (variates spanning greater intervals intercorrelate more strongly than those spanning lesser times as elaborated in

TABLE VII
INTERCORRELATION MATRICES FOR ITEM ANALYSES OF $p(R_2)$ UNDER TWO CONDITIONS OF RECALL[a]

| | | Retention interval | | |
|---|---|---|---|---|
| | | 20 Minutes | 2 Days | 28 Days |
| Not stimulated | | | | |
| | 2 Minutes | .58 | .55 | .62 |
| | 20 Minutes | — | .56 | .65 |
| | 2 Days | — | — | .59 |
| Stimulated | | | | |
| | 2 Minutes | .51 | .40 | .46 |
| | 20 Minutes | — | .42 | .44 |
| | 2 Days | — | — | .30 |

[a] Data from Bilodeau & Blick (1965).

Table I; see Jones, 1962) just as discussed at some length on pages 330–332 in the previous section. The source of this variance change is the stimuli introduced by the experimenter into the retention-test environment and cannot be attributed to repeated recalling since there was none. Another point well worth noting is that the amount of variance accounted for at a longer retention interval by reference to coefficients of previous days is larger than for shorter retention intervals. There is no support for a common view that less and less variance is explainable the longer the retention interval and there is no support for the view that as forgetting gets more pronounced there is less and less to talk about. Instead, there is a good deal of obtained structure to explain at the longer intervals.

Current work in the laboratory at Tulane continues to move recall upward or downward by large amounts. We are having good luck at facilitating the recall of a secondary word when the norms show it to be an associate of the primary (Fox & Bilodeau, in press). For example, *bitter* ($S_1$) evokes *sweet* ($R_1$), *sweet* evokes *sour*, and *bitter* evokes *sour* ($R_2$). *Sour* is well remembered when it is in one list and *sweet* in the other. Thus, when these interassociations obtain, primaries convert to secondaries in recall and great facilitation of secondaries is obtained instead of Blick's interference-by-primary phenomenon. In order to do research like this, it is necessary to have a large number of normative networks, and we have them from 1000 subjects free-associating three times (method of continued association) to 138 words (Bilodeau & Howell, 1965). These norms contain hundreds of matrices and many laws on associative processes and networks. Another line of inquiry involves assessing an individual's free-associative structure and then arranging the conditions of recall so as to move *his* recall upward or downward as desired. Here we are comparing the predicted and observed utterances of individual subjects. Fox (1964) has interpolated a period of free association between Training and Recall. The subjects in the experimental groups free associated to stimuli corresponding to the $R_2$s exposed as training words. There were strong trends in Recall depending upon the correspondence of the free associate and the $R_2$ of Training. In another one of our investigations, while trying to keep the gimmicks at a minimum, we hope to discover that there are conditions for which retention by stimulation is better after longer rests than shorter ones. Taken as a whole, the studies above all show a stress on associative processes and networks and are part of a rapidly expanding research interest in communication shown in the literature at large.

### IMPLICATIONS FOR MOTOR TASKS

It has occurred to the writer more than once that the stimulus is poorly identified in motor-skills articles. This would in turn handicap those who wish to work with associative strength or to relate responses to stimuli in any molecular and precise manner. The writer and his colleagues seem to have made some progress on this difficult problem by identifying $R_1$, $R^a$, and KR as the within-trial events to which recall events can be related in some more or less precise way. No basic distinction between motor and verbal events was made. Rather, the discussion turned on expected and obtained response associates of the training events. No S-R hierarchies of associative strength for levering responses existed prior to the experiments reported in the section on free recall and it was not possible to run word and levering analogs in any strict sense.

Melton (1964) believes that control-display (S-R) population stereotypes might be employed to test the interference theory of forgetting. For example, he infers, though his data have not yet supported the inference, that a subject taught a tracking task under conditions of low control-display preference will remember more poorly (than if trained under high preference) when the high preference shows recovery over rest. Fitts (1964) is the prominent figure in the field of stereotypes and his S-R compatible-incompatible model for switch-relationship preferences in space is one kind of associative hierarchy. The psychometric assessment of associative structures, except for word-word hierarchies, is in its infancy and should be given every consideration in future projects. Probably, a knowledge of several hierarchies will be necessary to explain much variance for any given circumstance in any complex skilled act. The distinction between the motor and verbal domain was always thought to involve considerable overlap and now the distinction is even less useful (Bilodeau & Levy, 1964; Fitts, 1964). Chapters 1, 2, and 6 on Selective Learning, Mechanisms of Response, and Information Feedback, respectively, emphasize the need to use S-R information from more than one source and show that the motor investigator must include more traditional verbal and perceptual stimuli and reactions in his prediction equation if he is to explain more of motor behavior. Alternatively, to explain response variance, several sources must be consulted, only one of which is motor.

## Summary

The stimulation technique is a broad and useful method for modifying tests of free recall in order to study the retrieval of stored information. The technique is not unrelated to other methods already in use. Basically, reminders have been or can be used during recall when the subject is again presented with stimulus, response, and feedback material associated in various degree with the events to be retrieved. Reminders have or can be used when the environment is arranged so that the stimuli and responses occurring during recall will themselves tend to call out those used during training. The past few years have seen a marked increase in interest in associative structures, their measurement and experimental employment. This is especially true of verbal learning and memory. It is not true of motor learning and memory where normative data on associative hierarchies and networks are lacking. Much work lies ahead in comparing free association, free recall, and stimulated recall.

Until now, S-R variations have been largely confined to periods of training and interpolated training because the fashion during the past 60 years has been to stress the learning process over the memory process. To exaggerate, we have often treated the recall period as if training events were always spewed out in the absence of recall stimulation. If we are to make use of a more complete S-R psychology, it will be necessary to regulate the S-R environment of the recall period as fully and as ingeniously as we have done for the acquisition period and then to examine the structure of the recall data. One way to begin, it was suggested here, is to consider the rules of free association as a cornerstone of memory and then proceed to elaborate the rules.

## References

Adams, J. A. The second facet of forgetting: a review of warmup decrement. *Psychol. Bull.*, 1961, **58**, 257–273.

Ammons, R. B., Farr, R. G., Bloch, E., Neumann, E., Dey, M., Marion, R., & Ammons, C. H. Long-term retention of perceptual-motor skills. *J. exp. Psychol.*, 1958, **55**, 318–328.

Bahrick, H. P. Retention curves: facts or artifacts. *Psychol. Bull.*, 1964, **61**, 188–194.

Bahrick, H. P. The ebb of retention. *Psychol. Rev.*, 1965, **72**, 60–73.

Bahrick, H. P., & Bahrick, P. O. A re-examination of the interrelations among measures of retention. *Quart. J. exp. Psychol.*, 1964, **16**, 318–324.

Barnes, J. M., & Underwood, B. J. 'Fate' of first-list associations in transfer theory. *J. exp. Psychol.*, 1959, **58**, 97–105.

Battig, W. F., Nagel, E. H., Voss, J. F., & Brogden, W. J. Transfer and retention of bidimensional compensatory tracking after extended practice. *Amer. J. Psychol.*, 1957, **70**, 75–80.

Bilodeau, E. A., & Bilodeau, I. McD. Variation of temporal intervals among critical events in five studies of knowledge of results. *J. exp. Psychol.*, 1958, **55**, 603–612.

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Bilodeau, E. A., & Blick, K. A. Courses of misrecall over long-term retention intervals as related to the strength of pre-experimental habits of word association. *Psychol. Rep.*, 1965, **16**, 1173–1192, Monogr. Supplement No. 6-V16.

Bilodeau, E. A., & Howell, D. C. *Free association norms.* U.S. Gov't. Print. Office, Washington, D.C., 1965.

Bilodeau, E. A., & Levy, C. M. Long-term memory as a function of retention time and other conditions of training and recall. *Psychol. Rev.*, 1964, **71**, 27–41.

Bilodeau, E. A., Sulzer, J. L., & Levy, C. M. Theory and data on the interrelationships of three factors of memory. *Psychol. Monogr.*, 1962, **76**, No. 20 (Whole No. 539).

Bilodeau, E. A., Levy, C. M., & Sulzer, J. L. Long-term retention under conditions of artificially induced recall of related events. *Percept. mot. Skills*, 1963, **16**, 875–910 (Monogr. Suppl. No. 8-V16). (a)

Bilodeau, E. A., Fox, P. W., & Blick, K. A. Stimulated verbal recall and analysis of sources of recall. *J. verb. Learn. verb. Behav.*, 1963, **2**, 422–428. (b)

Bilodeau, E. A., Jones, M. B., & Levy, C. M. Long-term memory as a function of retention time and repeated recalling. *J. exp. Psychol.*, 1964, **67**, 303–309.

Blick, K. A. Cultural primaries as a source of interference in short-term verbal retention. *J. exp. Psychol.*, 1965, **69**, 246–250.

Boswell, J. J., & Bilodeau, E. A. Short-term retention of a simple motor task as a function of interpolated activity. *Percept. mot. Skills*, 1964, **18**, 227–230.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. *J. gen. Psychol.*, 1953, **49**, 229–240.

Bousfield, W. A., & Puff, C. R. Clustering as a function of response dominance. *J. exp. Psychol.*, 1964, **67**, 76–79.

Broadbent, D. E. Flow of information within the organism. *J. verb. Learn. verb. Behav.*, 1963, **2**, 34–39.

Cofer, C. N. Comparison of word associations obtained by the methods of discrete single word and continued association. *Psychol. Rep.*, 1958, **4**, 507–510.

Cofer, C. N. On some factors in the organizational characteristics of free recall. *Amer. Psychologist*, 1965, **20**, 261–272.

Coleman, E. B. The association hierarchy as an indicator of extraexperimental interference. *J. verb. Learn. verb. Behav.*, 1963, **2**, 417–421.

Conrad, R., & Hille, B. A. The decay theory of immediate memory and paced recall. *Canad. J. Psychol.*, 1958, **12**, 1–6.

Deese, J. On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. exp. Psychol.*, 1959, **58**, 17–22. (a)

Deese, J. Influence of inter-item associative strength upon immediate free recall. *Psychol. Rep.*, 1959, **5**, 305–312. (b)

Deese, J. From the isolated verbal unit to connected discourse. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior.* New York: McGraw-Hill, 1961. Pp. 11–31.

Ebbinghaus, H. *Uber das Gedächtnis: Untersuchungen zur experimentellen Psychologie.* Leipzig: Duncker & Humblot, 1885; Translated by H. A. Ruger and C. E. Bussenius as *Memory: a contribution to experimental psychology.* New York: Teachers College, Columbia Univer., 1913.

Fitts, P. M. Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning.* New York: Academic Press, 1964. Pp. 243–285.

Fleishman, E. A. A factor analysis of intra-task performance on two psychomotor tests. *Psychometrika*, 1953, **18**, 45–55.

Fleishman, E. A., & Parker, J. F., Jr. Factors in the retention and relearning of perceptual-motor skill. *J. exp. Psychol.*, 1962, **64**, 215–226.

Fox, P. W. Stimulated recall as a function of interpolated word associations. Unpublished doctoral dissertation, Tulane Univer., 1964.

Fox, P. W., & Bilodeau, E. A. Associative rules governing recall and misrecall. *J. exp. Psychol.*, (in press).

Fox, P. W., Blick, K. A., & Bilodeau, E. A. Stimulation and prediction of verbal recall and misrecall. *J. exp. Psychol.*, 1964, **68**, 321–322.

Garskof, B. E., & Houston, J. P. Supplementary note: Recall as a function of word association strength and associative overlap. *Psychol. Rep.*, 1964, **14**, 185–186.

Hunter, W. S. Learning, IV: Experimental studies of learning. In C. A. Murchison (Ed.), *A handbook of general experimental psychology*. Worcester, Mass.: Clark Univer. Press, 1934, Pp. 497–570.

Irion, A. L. The relation of 'set' to retention. *Psychol. Rev.*, 1948, **55**, 336–341.

Jenkins, J. J., & Russell, W. A. Associative clustering during recall. *J. abnorm. soc. Psychol.*, 1952, **47**, 818–821.

Jones, M. B. Practice as a process of simplification. *Psychol. Rev.*, 1962, **69**, 274–294.

Keppel, G., & Underwood, B. J. Proactive inhibition in short-term retention of single items. *J. verb. Learn. verb. Behav.*, 1962, **1**, 153–161.

Lewis, D., McAllister, D. E., & Adams, J. A. Facilitation and interference in performance on the Modified Mashburn Apparatus, I: the effects of varying the amount of original learning. *J. exp. Psychol.*, 1951, **41**, 247–260.

Lloyd, K. E., & Johnston, W. A. Short-term retention as a function of contextual constraint. *J. exp. Psychol.*, 1963, **65**, 460–467.

Luh, C. W. The conditions of retention. *Psychol. Monogr.*, 1922, **31**, No. 3 (Whole No. 142).

McGeoch, J. A. The comparative retention values of a maze habit, of nonsense syllables, and of rational learning. *J. exp. Psychol.*, 1932, **15**, 662–680.

Melton, A. W. Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.*, 1963, **2**, 1–21.

Melton, A. W. Retention of tracking skills. *Final Rep.*, Contract No. DA-49-007-MD-1020, U.S. Army Med. Res. and Develpm. Command and Univer. of Michigan (1964).

Mengelkoch, R. F., Adams, J. A., & Gainer, C. A. The forgetting of instrument flying skills as a function of the level of initial proficiency. *NAVTRADEVCEN tech. Rep.*, 1958, No. 71-16-18.

Murdock, B. B., Jr. The immediate retention of unrelated words. *J. exp. Psychol.*, 1960, **60**, 222–234.

Murdock, B. B., Jr. Interpolated recall in short-term memory. *J. exp. Psychol.*, 1963, **66**, 525–532.

Naylor, J. C., & Briggs, G. E. Long-term retention of learned skills: A review of the literature. *USAF ASD tech. Rep.*, 1961, No. 61-390.

Noble, C. E. An analysis of meaning. *Psychol. Rev.*, 1952, **59**, 421–430.

Peterson, L. R., & Peterson, M. J. Short-term retention of individual verbal items. *J. exp. Psychol.*, 1959, **58**, 193–198.

Postman, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw-Hill, 1961. Pp. 152–179. (a)

Postman, L. Extra-experimental interference and the retention of words. *J. exp. Psychol.*, 1961, **61**, 97–110. (b)

Postman, L. Short-term memory and incidental learning. In A. W. Melton (Ed.), *Categories of human learning*. New York: Academic Press, 1964. Pp. 145–201.

Reid, L. S., Brackett, H. R., & Johnson, R. B. The influence of relationships among items to be recalled upon short-term retention. *J. verb. Learn. verb. Behav.*, 1963, **2**, 86–92.

Reynolds, B. The effect of learning on the predictability of psychomotor performance. *J. exp. Psychol.*, 1952, **44**, 189–198.

Reynolds, B., & Bilodeau, I. McD. Acquisition and retention of three psychomotor tests as a function of distribution of practice during acquisition. *J. exp. Psychol.*, 1952, **44**, 19–26.

Roehrig, W. C. Psychomotor task with perfect recall after fifty weeks of no practice. *Percept. mot. Skills*, 1964, **19**, 547–550.

Rothkopf, E. Z., & Coke, E. U. The prediction of free recall from word association measures. *J. exp. Psychol.*, 1961, **62**, 433–438.

Russell, W. A., & Jenkins, J. J. The complete Minnesota norms for responses to 100 words from the Kent-Rosanoff Word Association Test. *Tech. Rep.* No. 11, 1954, Contract No. N8onr-66216, Office of Nav. Res. and Univer. of Minnesota.

Shepard, R. N., & Chang, J. J. Forced-choice tests of recognition memory under steady-state conditions. *J. verb. Learn. verb. Behav.*, 1963, **2**, 93–101.

Tsai, L. S. The relation of retention to the distribution of relearning. *J. exp. Psychol.*, 1927, **10**, 30–39.

Tulving, E. Intratrial and intertrial retention: notes towards a theory of free recall verbal learning. *Psychol. Rev.*, 1964, **71**, 219–237.

Underwood, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49–60.

Underwood, B. J., & Postman, L. Extraexperimental sources of interference in forgetting. *Psychol. Rev.* 1960, **67**, 73–95.

Yntema, D. B. Keeping track of several things at once. *Hum. Factors*, 1963, **5**, 7–17.

# Methods of Measuring Retention[1]

## Comments on Professor E. A. Bilodeau's Paper

HARRY P. BAHRICK
*Ohio Wesleyan University*

The methods of measuring retention remained largely unchanged from the earliest days of experimental psychology to the 1950's. We are indebted to Bilodeau for his many original contributions to the rebirth of interest in this central problem. We are also indebted to him for the present discussion in which he reviews the significant work which has been done in this field during the past decade. He raises a number of issues, the discussion of which is certain to stimulate further progress in this area.

As a general preliminary issue it appears to me that the criteria for evaluating the methods of the experimental psychologist require a brief discussion. How do we decide whether a method is useful, or whether it is better than some other method? This problem was first discussed by Melton (1936), later by Fitts, Bahrick, Noble, and Briggs (1959), and most recently by Bahrick and Noble (1966). Reliability and validity are, of course, the principal demands we make of our measures, but these characteristics need a redefinition which is relevant to the objectives of the experimental psychologist. We demand reliability only in the sense of a "noise-free" measurement, and this does not necessarily mean a measurement which remains stabile for a particular individual. On the contrary, it is quite possible that insensitive scores will remain stabile while more sensitive indicants will fluctuate and thereby reveal more adequately the phenomena we wish to assess. The identification of various sources of instability of an indicant is frequently difficult, but becomes an important requirement in exploiting its information potential.

---

[1] The ideas expressed in this paper were developed in the performance of work supported by Public Health Research Grant HD 00926-04 from the National Institute of Child Health and Human Development.

As to validity, this still relates to the question of whether we are measuring what we intend to measure, but no simple test-criterion correlation can provide the answer. Rather we must judge the validity of our methods by the clarity of the answers they yield to the questions which our theories raise. An experimental method is valid if it fulfills its purpose, i.e., if it yields an unequivocal answer to the question the experimenter has asked. If four different indicants are used to reveal what has been learned during a conditioning experiment, all of them have *some* validity to the extent that they show systematic behavior changes during the conditioning process. If the intercorrelations among them are low, however (as it happens they are), then the measures are either unreliable, i.e., they reflect random uncontrolled variables, or they systematically reflect some differential aspects of the changes which occur during conditioning. It is my contention that such diversity among indicants of the phenomena that interest us is not a liability, but a potential asset. If an architect draws three different views of a building, each view provides valuable additional information, but the information can be exploited only to the extent that we are able to interpret the three views differentially. Obviously the additional information provided by each indicant cannot be exploited by treating the indicants interchangeably, i.e., ignoring their demonstrated differences and assuming that each one of them reveals equally and completely what we wish to find out. We cannot exploit the potential validity of an indicant until we know more precisely what it measures that the other indicants do not measure. To find this out we require rational analysis and empirical observations of their interrelations as a function of a variety of independent variables. It is this kind of study, of the interrelations among response variables, that I believe has been greatly neglected in many areas of experimental psychology and has limited the validity of our methodology. Two conclusions are implicit if the above considerations are applied to the measurement of retention. "The more methods the merrier" is the first conclusion since each new method is an addition to our potential sources of information. The second conclusion is the need for intensive analysis of the interrelations among new and old methods. Prior to obtaining the results of such an analysis we must exercise due restraint about statements concerning memory in general, or about the shape of *the* curve of retention. This is true because any such observations will reflect to an unknown degree characteristic interactions of memory and method, i.e., artifacts of measurement.

Viewed from this perspective Bilodeau's work represents a most significant contribution to our study of memory. He has added several new and analytic indicants to augment the recently reawakened interest in methodology in an area which had been dormant for several decades. By interpreting these, however, he has opened a Pandora's box of new problems and has stuck out his neck in a way which seems hazardous when judged by the standards of caution which I have recommended for others—but which rarely inhibit my own pronouncements.

Let us first examine Part I of his chapter. Here Bilodeau has pointed out the important fact that the set of events which we call a trial involves relatively easily separable subevents, such as the response itself and the experimenter feedback. He has further shown that independent measurement of the retention of these subevents, and of their interrelations, can improve our understanding of the phenomena of retention. The component events of a trial had either been ignored or confounded in the past. In sorting these out and empirically examining their interrelations he has made a badly needed contribution to more analytic measurement. Perhaps Bilodeau's success in destroying the pseudo-unit of a trial will lead us to a more critical reexamination of this concept in other contexts, such as the testing of all-or-none models of one-trial learning.

In his discussion of measures of memory Bilodeau begins by pointing out the central fact that memory curves usually record the decline of performance of a group mean, and that this measure leaves out much information concerning the changes of a whole distribution of responses in time. No one can quarrel with this fact, nor with the inference that follows, namely, that we need far more complete and analytic information about changes in the distributions of responses produced by subjects at various points in time. As a partial solution Bilodeau has suggested an examination of correlations among the distributions as well as an examination of changes in their variance. I find both of these suggestions intriguing and the data he reports significant, but I have strong reservations concerning Bilodeau's interpretation of these measures. It is here where I believe that he has stuck his neck out very far and I propose to reexamine these interpretations.

Let us consider first the interpretation of the proposed correlation measure of retention. In this method, subjects are instructed to reproduce a previously made motor response. The first and second array of responses produced by the subjects are correlated, and it is shown that the correlation decreases as a function of the time which intervenes between the two events. I have several reservations about interpreting the decrease in correlation as an indicant of retention, and I have even greater reservations concerning the comparison of relative amounts of forgetting over successive time intervals on the basis of the amount of decline of the correlation function. In other words, I am not certain that retention is being measured and I question, even more, the implication that it is being measured with an equal-interval scale.

My principal objection to interpreting the correlation function as a measure of retention arises from the lack of specification of exactly what has been learned. If we conceive of retention as the persistence of the changes which occurred during learning, then we clearly must first establish what these changes are before we can decide to what extent they have persisted in time. Psychologists have advocated caution in the past about the conclusion that a loss of retention occurred if the previous learning process could not be established beyond question. The senile patient, for example, who cannot report what he ate for breakfast may be described as forgetful by many people, but the psychologist has insisted on the significance of differentiating the failure to retain the traces left by an experience from the failure of the experience to produce traces in the first place. The mere presentation of a stimulus, or the occurrence of a response, is certainly not adequate evidence of learning. Rather, it has generally been held that learning must be established on the basis of some *change* of performance and the minimal requirement for the assessment of such a change is a comparison of two measures of performance during practice. No such two measures are available for the correlation index. The array of individual differences among the responses of the various subjects is not a distribution which has been established as the result of practice; rather it is an array which has been observed on the occasion of the first response the subjects were instructed to make (Bilodeau & Levy, 1964; Bilodeau, Sulzer, & Levy, 1962). It is, therefore, not possible to establish any relations between changes in the array which are observed later on and changes which occurred during training. If this criterion is not met for measuring retention, then all changes which occur in time in a distribution of responses produced by many individuals can be called forgetting. Clearly this writer does not agree with Bilodeau's definition of memory as persistence, and of forgetting as inconsistence. We must ask: Persistence of what? Only if we can establish that the persistence relates to changes that were systematically established during training does the later observation of inconsistence preserve the traditional meaning of the concept of forgetting. Bilodeau could counter with the claim that his subjects were instructed to reproduce their original responses as nearly as possible, and that the loss of the correlation in time provides an index of their decreasing ability to follow these instructions. I wonder, however, whether a very similar phenomenon might not be observed if the subjects either (a) were instructed to make another response, i.e., without being specifically told to reproduce their original response, or (b) were given no error information in connection with their first response, and on the occasion of their second response they were again told to produce a response which they imagined to be of the desired type, rather than to attempt to reproduce their first response. Bilodeau has pointed out that at least the early loss of correlation is probably due to a loss of set which occurs when subjects briefly walk away from the apparatus and return before attempting to reproduce their first response. When subjects are confronted twice with the requirements to make a response under identical stimulus situations they tend to make similar responses for a variety of reasons, among them a general set. If more time intervenes between the two occasions to respond, this becomes less true. I do not believe that this lessened tendency for response consistency in time can be interpreted unequivocably as an indicant of retention. At the very minimum we would need to contrast the loss of correlation for the group instructed to reproduce their original responses with the correlation loss of control groups given the original response instructions, without the additional request to reproduce their first response.

Let me comment here briefly on why this type of control is not equally needed in many other studies of retention. In most studies of verbal learning, for example, such controls are not necessary because we are dealing with performance which has a very low *a priori* probability of occurrence. A subject is quite unlikely to come up with a correct sequence of nonsense syllables, words, or numbers during a retention test, if he did not learn them during the practice trials. For this reason the type of control I have described is unimportant also in most short-term retention studies, despite the fact that they involve only a single acquisition trial. One is usually justified in assuming that the correct responses on a retention test reflect the effect of this prior learning trial. The situation is different, however, in a positioning task in which the subject rather than the experimenter has produced the response to be remembered, and in which *a priori* response sets may contribute substantially to response consistency.

It seems very likely to me that response sets or response consistency is greater for certain subjects than for others. This tendency alone would cause the correlation of the arrays of scores to decrease in time, quite aside from the process that we usually describe as retention. The correlation measure is simply an index of the degree of persistence of individual differences observed to occur on the occasion of a first response. Such progressive changes in individual differences can occur for a variety of reasons and cannot, in my opinion, be interpreted as measures of retention, even if subjects performed the second response under instructions to attempt to reproduce the first one. Admittedly, the meaning of the term *retention* is to some extent arbitrary, like the meaning of all words, but there appears to me to be considerable advantage in reserving the term for the persistence of those changes in performance which were observed to occur during practice and not to apply it more generally to the presistence of distributions of individual differences which happened to be observed on the occasion of a first response in a task.

In his analysis of the shape of retention curves (p. 327) Bilodeau contrasts the type of curve he obtained from the decline in the correlation measure with the negatively accelerated curve most frequently reported for other measures. I have described elsewhere (Bahrick, 1965) my reservations about the interpretation of slopes of retention curves based upon classical measures. The reservations are based upon the fact that the traditional indicants confound several variables and I shall not review these arguments here. However, the problem of confounding of variables applies probably to an even greater degree to the curves analyzed by Bilodeau. As was suggested above there are many variables, not yet individually assessable, which may contribute to the instability of a distribution of individual differences in time, and the rate of loss of correlation reflects all of them, and possibly to a different degree at different times. Further, the correlation measure itself does not give a linear index of variance which is predictable in the second distribution from a knowledge of the first distribution. The square of the coefficient (the coef-
which can be accounted for from a knowledge of the predictor scores. It would seem to me that this coefficient would be a more appropriate indicant of consistency between the two distributions of responses than the correlation coefficient itself. The shape of the retention function would, of course, differ markedly depending upon whether $r$ or $r^2$ is plotted; and the arbitrary selec-
tion of $r$ provides a further basis for my strong reservations about interpreting the resulting curve as an equal-interval indicant of how much is forgotten. In summary, I believe that the gradual change in time of a distribution of indi-
vidual differences, which Bilodeau reports in the form of a correlation func-
tion, is an interesting phenomenon. It should be related to other independent and dependent variables, but should not at the present time be interpreted as an index of retention and most certainly not as an equal-interval indicant of retention.

Let us now turn our attention to the proposed variance measure of retention. To obtain this measure subjects are instructed to reproduce a response which they previously made, and the variance of the response distribution they generate is plotted as a function of the time intervening between the original and the reproduced response.

My principal objection to the interpretation of the correlation measure applies equally to the interpretation of the variance measure, as used by Bilodeau and Levy (1964), but I feel much friendlier toward this measure because the basis for my objection could easily be remedied. The objection, of course, hinges on the fact that Bilodeau and Levy do not report how variance of the response distribution changes during training and, therefore, I am unwilling to accept *a priori* the progressive variance changes of later response distributions as indicants of forgetting. It would seem a simple matter, however, to establish the variance of successive response distributions during training and to plot a "learning curve" on this basis. One would certainly expect that a group of subjects will diminish the variance around a desired response as practice of the response continues with knowledge of results. Subsequent increases of variance could then be established as indicants of forgetting. An analogous validation of the correlation measure would yield less predictable results. It would be necessary to correlate the individual difference distribution obtained on a final practice trial with the individual difference distribution obtained on each of the preceding trials, and to plot the increase in the correlation coefficient as a function of practice. Presumably, the highest correlation would be obtained between the final practice trial and the immediately preceding one, as suggested by the superdiagonal form shown in Bilodeau's Table I. It would surprise me, however, if such a correlation would ever approach unity, since it is limited not only by "noise" factors, but also by differential rates of learning among subjects which would alter successive distributions of individual differences. Even a high correlation obtained on this basis reflects in part the persistence of a set between trials and its loss provides a questionable basis for the measurement of retention. In any case, a retention curve based on the correlation measure would have to use the observed maximum correlation during training as the first ordinate value of the retention curve. It seems to me that this would rule out Stage I of Bilodeau's curve in Fig. 10, which assumes a value of unity on the ordinate at the end of training.

The use of the variance measure on the other hand is unencumbered by problems of nonlinearity of the indicant; a variance value can easily be established at the end of training and there is ample precedent for the use of variance as a measure of learning. The traditional practice has been to compute variance among successive responses of the *same* subjects, and then determine the average variance scores for a group of subjects. Variations of the psychophysical method of average error illustrate this technique for discrete tasks (Bahrick & Noble, 1961), and the root mean square error is an analogous indicant of performance in continuous tasks. The variance measure of Bilodeau and Levy differs from these principally because it is computed for a single response made by many subjects, rather than for many successive responses made by a single subject. Bilodeau and Levy's technique is more promising for the measurement of retention because it avoids the problem of confounding the retention measure with further practice in a task—a problem which cannot be avoided if a series of successive responses of the same subject is needed. By establishing the relations of this variance measure with other measures of retention as a function of a variety of procedural and task variables, we will undoubtedly gain considerable insight into the nature of learning and forgetting. Until this is done I believe that discussion of the shape of a general retention curve based upon the variance and correlation indicants is premature. Bilodeau's analysis of the curve into a two-phase process is intriguing but based upon too many methodological limitations to be regarded as more than an interesting hypothesis. To sum it all up, I do not believe that we can make meaningful interpretations of the slope of retention curves as long as we are uncertain about the availability of equal-interval scaling on the ordinate of our curves, and know little or nothing about the relation of the units we plot for one indicant to the units of other indicants of retention.

I am greatly impressed with the significance of the work on stimulated recall which Bilodeau describes in the second part of his chapter. There can be no doubt that memory generally involves complex networks of associations. These associations occur in part because our experiences are embedded in a stream of consciousness which brings about the bonds of temporal contiguity. But the associations also occur because a wealth of knowledge established in the past can mediate concepts which bridge the gaps of time and space. Thus, if we try to recall where we have placed a lost object we may first mentally retrace our steps so as to take advantage of the associations established by temporal contiguity. If this fails to retrieve the memory, however, we are likely to make extensive use of other associations relating to our past habits with the class of object involved.

Most laboratory work on retention has failed to come to grips with this complex network of associations which generally provides the basis for retrieving stored information in everyday life. In the quest for increased control, investigators have concentrated most of their efforts upon meaningless material, i.e., material which must be retrieved principally on the basis of the temporal associations formed during the actual experimental sessions. The resulting models of retention have reflected this comparative lack of involvement of past experience. I have no doubt that a probabilistic concept of memory based upon multiple associations of both the temporal and the mediated type will bring our models closer to functional reality. I see the work on stimulated recall, described by Bilodeau, Fox, and Blick (1963), as a significant step in this direction. It should yield a family of retention curves far more diversified in slope than the curves we have so far observed in our laboratories. In turn this will help our models to account more adequately for associations which we remember throughout most of our life, rather than only for those which are largely lost within a few hours after learning. Beyond this, the stimulated recall situation may yield several means of scaling degrees of associative strength below the unaided threshold of recall. This should contribute to the solution of methodological and theoretical problems. By making a scale of associative strength available below the unaided recall level, we may avoid artifacts of dichotomous measurement and also obtain more definitive answers to questions arising from the conflict of discrete and continuous models of acquisition and retention.

Another potential contribution of the method of stimulated recall lies in the possibilities this method offers for reaching an empirically founded rapprochement between dynamic theories of forgetting and the more atomistic interference theories. The dynamic theories have emphasized long-range changes of the content of recall, based upon the context of long established trace systems. The S-R theories have concentrated upon short-term losses of retention, more easily measured in the laboratory. Until recently (Underwood, 1957), S-R theories have attributed the loss of retention primarily to the effects of retroactive inhibition. By using the method of stimulated recall and by taking prior measures of existing associative systems, Bilodeau, Fox, and Blick were able to extend predictions to the nature of changes of subject's responses. This work suggests the extension of controlled laboratory techniques to cover the more elusive and long-range phenomena which until recently were dealt with only by the less researched dynamic theories.

Bilodeau has stated the point very clearly (p. 345): There is no support for a common view that less and less variance is explainable the longer the retention interval, and there is no support for the view that as forgetting gets more pronounced there is less and less to talk about.

It would seem that the extension of laboratory methods of measuring retention, which Bilodeau has so ably summarized, will bring us closer to a realistic involvement with our ultimate subject matter: human behavior.

## REFERENCES

Bahrick, H. P. The ebb of retention. *Psychol. Rev.*, 1965, **72**, 60–73.

Bahrick, H. P., & Noble, M. E. On stimulus and response discriminability. *J. exp. Psychol.*, 1961, **61**, 449–454.

Bahrick, H. P., & Noble, M. E. Motor behavior. In J. B. Sidowski (Ed.), *Experimental methods and instrumentation in psychology.* New York: McGraw-Hill, 1966 (in press).

Bilodeau, E. A., & Levy, C. M. Long-term memory as a function of retention time and other conditions of training and recall. *Psychol. Rev.*, 1964, **71**, 27–41.

Bilodeau, E. A., Sulzer, J. L., & Levy, C. M. Theory and data on the interrelationships of three factors of memory. *Psychol. Monogr.*, 1962, **76**, No. 20 (Whole No. 539).

Bilodeau, E. A., Fox, P. W., & Blick, K. A. Stimulated verbal recall and analysis of sources of recall. *J. verb. Learn. verb. Behav.*, 1963, **2**, 422–428.

Fitts, P. M., Bahrick, H. P., Noble, M. E., & Briggs, G. E. Skilled performance. Ohio State Univer. Res. Found., Final Report, Contract AF 41(657)-70, March 1959.

Melton, A. W. The methodology of experimental studies of human learning and retention, I: the functions of a methodology and the available criteria for evaluating different experimental methods. *Psychol. Bull.*, 1936, **33**, 305–394.

Underwood, B. J. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49–60.

## Chapter 8

# Tracking Behavior

E. C. POULTON
*Medical Research Council, Applied Psychology Research Unit, Cambridge, England*

## Introduction

### Five Kinds of Tracking

*Pursuit* tracking typically involves keeping a marker in line with a moving target. An example is the pursuit rotor, in which the tip of a stylus has to be held on a small circular disk set into a rotating platform (Ammons, 1955).

In *compensatory* tracking there is only one moving element. This has to be held stationary over a fixed target, from which it tends to move away. In both these cases movement is present in the display in the absence of a response.

In *acquisition* or discontinuous *step-function* tracking, the target and response marker typically start superimposed. One of them then suddenly jumps to a new position, and the man has to superimpose the marker on the target again. It has always been assumed that the tasks are equivalent whichever of the two, target or response marker, jumps to the new position. These two acquisition tasks may be regarded as special cases of pursuit or compensatory tracking in which, except for the sudden jumps, movement is only present in the display as a result of the man's response. In most practical situations acquisition precedes pursuit or compensatory tracking; the target has to be acquired before it can be held in alignment.

The two remaining kinds of tracking have received much less study. Both involve a target extended in space like a wiggly line. By far the commonest in everyday life can be called *unpaced contour* tracking. Here the man moves along the contour at his own speed, like the driver of a car going along a wiggly road. Laboratory examples are the stylus maze, and the star pattern often traced in mirror-drawing experiments.

In *paced contour* tracking the wiggly line is normally drawn on a paper tape which approaches the man at a fixed speed. He controls a marker which moves across the paper tape at right angles to its direction of approach. He has to keep the marker superimposed upon the wiggly line as it passes (Poulton, 1954). In both kinds of contour tracking the man can normally see the wiggly line some distance ahead. When this preview is taken away, the task becomes more like pursuit tracking.

## Characteristics of Tracking Tasks

### *Precision of Movement*

All behavior can be said to involve deciding what to do and then doing it. As we have seen in Chapter 2 of this volume, the study of selective motor learning is concerned mainly with deciding what to do. For example, the man may have to decide which one of four buttons to press, or in which of four directions to move a lever. The excellence of his performance is assessed in terms of time taken and whether the response selected is the correct one or not. The button or lever has simply to be moved sufficiently far to make an electrical contact. The exact distance it moves, and the exact force exacted on it by the man, are not considered.

In contrast, the study of tracking is concerned more with the execution of the response. Here performance is assessed in terms of the distance moved or the force exerted. Typically the man is supplied with only one control; he does not have to choose between controls as he often does in selective motor learning. Thus selective motor learning and tracking emphasize different aspects of motor behavior.

At the present time the study of tracking is in a sense the more unique, in that it has less in common with the study of verbal learning. Like the study of selective motor learning, the study of verbal learning is largely confined to response selection. However, there is no need for this, and in the future considerably more emphasis in verbal learning will doubtless be placed upon the execution of responses. This is a problem which has always faced the man learning to speak a foreign language. But it has now become the problem also of the man designing the input to the computer which will in time automate the job of the shorthand typist. The future executive wishing to use such a machine will have to learn a pronunciation acceptable to the input device. Thus, whereas in the field of response selection the study of motor learning has drawn heavily upon the corresponding studies of verbal learning, in the field of response execution the roles may well be reversed.

### *Pacing*

The typical pursuit or compensatory tracking task is paced. At any instant the man has a movement to make of a specific size, but the size changes all the time. Making the correct size of movement either too early or too late is equivalent to making the wrong size of movement, and is penalized accordingly. By being paced the typical tracking task again differs from selective motor learning, since here the task may be either paced or unpaced, depending upon the wishes of the experimenter.

There are, however, unpaced tracking tasks. The principal one is target acquisition, or discontinuous step-function tracking. Here the size of movement required remains unchanged, so that its exact timing is not critical. In addition, contour tracking may be unpaced as well as paced.

### *Engineering Variables*

A difficulty in discussing tracking behavior is that there are many different tracking tasks, and that it is not possible to state categorically which task is the most typical. In studying tracking the experimenter has to decide what kind of track he is going to use, what kind of display, what kind of control, and what kind of control-system dynamics. In each case he has to make a more or less arbitrary selection from a number of possible alternatives, and his experimental results will be determined by his choice. Some of these independent variables have their counterparts in selective motor learning, but not all do. Since we shall be attempting to give as simple a picture of tracking behavior as is compatible with the facts, we shall assume that all independent variables other than the ones under discussion have "neutral" values; this means values which complicate the picture as little as possible. Clearly the independent variables must interact, but the effects of the isolated variables must be understood before we look at the interactions.

The control-system dynamics may be so complex that they completely dominate the man's performance. If this is so the tracking study is of little interest to the psychologist, although it may be the bread and butter of a design engineer concerned with a new vehicle such as a nuclear submarine or spacecraft. The complexity of the dynamics forms a kind of continuum, from the simple control systems which are studied principally by the psychologist, to the highly complex control systems which are studied principally by the engineer. Unless otherwise indicated, we shall restrict ourselves to the simplest control-system dynamics: a positional control without a control time lag or an unusual control gain. For most of the work of interest to psychologists has been carried out like this.

## The Track

### STEP FUNCTIONS

#### *Isolated Steps*

Step-function tracking or simple acquisition has the advantage that the visual stimulus to which the man responds can be defined precisely both in extent and in time. Usually the man is told to restore the alignment as quickly as possible. When he cannot anticipate the stimulus, his reaction time is almost independent of the size of the step. With a positional control his rate of movement and acceleration vary with the amplitude of the step, being greater for larger steps (Searle & Taylor, 1948; Taylor & Birmingham, 1948). This first shot at restoring alignment tends to be accurate to within about ±10 percent (Vince, 1948b, Fig. 4). If an appreciable error remains and he has plenty of time, he may have another shot, which will also be accurate to within about ±10 percent (Craik, 1947).

It seems clear from the consistent pattern of acceleration of the movement (Peters & Wenborne, 1936; Taylor & Birmingham, 1948), and from the short time which it normally takes (Craik, 1947; Taylor & Birmingham, 1948), that the movement is carried out without voluntary control. If the experimenter unexpectedly alters the apparatus and thereby tricks the man into making the wrong movement, it takes him almost .3 second after he notices that something is wrong before he can correct it (Hick, 1949). By this time the initial incorrect movement will normally have been completed (Taylor & Birmingham, 1949).

#### *Pairs of Steps*

When the stimulus for a second step is presented soon after a previous stimulus, the second response is likely to be delayed (Vince, 1948a, Exp. 4; Ellison & Hill, 1948). Collateral evidence suggests that this may be partly because the man is not expecting a second stimulus so soon, and so is not ready for it (Mowrer, 1940; Poulton, 1950a). But if the interval between the two stimuli is less than about .2 second, the response to the second signal will be delayed by about .1 second even when the man is half expecting to have to make a second response at this time (Poulton, 1950a; Davis, 1956).

When the man is preparing to deal with the expected second stimulus as well as with the first, his response to the first stimulus may be delayed by about .04 second (Gottsdanker, Broadbent, & Van Sant, 1963). It is as if the man reserved some of his computing capacity in readiness to receive the second stimulus, and so had inadequate capacity available to deal at maximum rate with the first stimulus. When the second step follows the first by an interval as short as .05 second, the man may on occasions be able to make a single response to the two steps (Craik, 1948).

Step-function tracking has not proved a very successful method of studying this so-called "psychological refractoriness." The two responses tend to run into each other when the interval between the stimuli is short, and it is not always possible to tell where the first response ends and the second response begins. Thus while Craik (1948) used the data to support a theory of two independent responses, or of a single response to the two stimuli, Ellson and Hill (1948) were able to use almost identical data to support a theory of the second response summing algebraically with the last part of the first response. The difficulty of interpretation can be overcome fully only by forsaking step-function tracking and using instead two separate responding members, e.g., two fingers, as was done by Davis (1956) and by Gottsdanker, Broadbent, and Van Sant (1963).

### *Series of Steps*

When the man is presented with a random series of steps, he learns the statistical properties of the series, and makes use of this in responding—the so-called "range effect." Large steps tend to be undershot, while small steps tend to be overshot (Ellson & Wheeler, 1949; Slack, 1953a), as if the man's response were determined partly by the average size of step, and only partly by the stimulus. Similarly the man tends to move more quickly with small steps than he would normally, and more slowly with large steps (Craig, 1949), as if his rate of movement also were determined partly by the average size of step. The range effect tends to be shown in more marked form when the steps follow each other in rapid succession (Slack, 1953b).

If the man is presented with a repeating series of steps, he soon learns it. He may then initiate a movement a little too soon, like a premature reaction in the classical reaction-time experiment. Even if he does not do this, he can sometimes be caught out by a surprise step which calls for a movement of unexpected timing, direction, or size. The man may make the response appropriate to the expected step, instead of to the surprise step (Slack, 1953a). Responses may also be more accurate in amplitude when the sequence is known in advance (Poulton, 1956).

## RAMP FUNCTIONS

The ramp function used most frequently in tracking has a constant velocity, but constant acceleration and constant rate of change of acceleration ("jerk") tracks have also been used. In the early stages of practice with a constant-velocity track, a compensatory display, and a positional control, Craik (1947) noted that the man corrected the error on the average every .5 second, the range extending from .25 to 1.00 second. But soon these intermittent corrections had superimposed upon them a constant rate of movement. The man, as it were, produced a rate of response which more or less matched the rate of the track, and then intermittently corrected the error between the two. The rate element in the man's response was clearly apparent at the end of the ramp. He continued to respond at the same rate for a reaction time, until he realized that the ramp had ceased (Craik, 1947, p. 58; Vince, 1948a, Fig. 1C).

A rather similar picture was described by Ellson, Hill, and Gray (1947) after practice with a pursuit display. The tracking records showed lengths lasting up to 2.0–4.0 seconds where the rate of the man's response function more or less matched the rate of the track, and since the misalignment was small, the man made little or no effort to correct it. These periods of pure "rate tracking" were shorter with the faster tracks, because a serious misalignment was likely to develop more rapidly. When a misalignment developed, it was corrected discretely at an average rate of about twice per second, the distribution of times corresponding to that described by Craik. Since the periodicity of the corrections was more or less independent of the rate of the track, the mean error of misalignment and the mean size of the corrections increased as the rate of the track increased.

Garvey and Mitnick (1957) described the course of learning a constant-rate track with a compensatory display and a positional control. Early in practice the man tends to lag behind the track, especially when the track is a fast one. But after practice the time lag is reduced to negligible proportions. The proportional reduction in the mean error of misalignment is of course a good deal less, since there is a residual oscillatory error even when the man, on the average, is on time. The man can also learn a constant acceleration track, but not a constant rate of change of acceleration ("jerk") track (Garvey, 1960).

When a quasi-random series of ramps of constant slope has to be tracked, the man can make use of his knowledge of the statistical properties of the series (Poulton, 1957b). This corresponds to the "range effect" found with step functions. The man can also learn a repeating series, and use this information in his tracking. A quasi-random series of ramps of constant slope can be tracked rather more accurately than a quasi-random sinusoidal track of similar amplitude and number of reversals per minute (Heinemann, 1961, problems A & D).

## Sinusoids

### *Single Sine Waves*

The task set in tracking a single sine wave depends upon its frequency. When the frequency is only 1 or 2 cycles per minute or less, the task is not unlike that of tracking a ramp function. After practice the man should be able to match the rate of the track much of the time. The only additional information which he needs to keep in mind, not required with ramp functions, is where he is on the sine wave, so that he knows whether the track is gradually accelerating or decelerating.

In contrast, when the frequency of a single sine wave is much above 60 cycles per minute, the man has not got time to respond separately to each half cycle, since his comfortable rate of responding averages only about 2 responses per second. If he responds discretely, his response has to include a reversal in the direction of his control movement, so that a movement in one direction merges into a return movement. The oscillographic charts from a practiced man suggest that the to-and-fro movements follow each other without a break. The man, as it were, generates a function something like a sine wave of the correct frequency and amplitude, and adjusts it from time to time to match the track. This is, of course, very much easier to do with a pursuit display than with a compensatory display, since the man can see the track he is trying to copy, and also the nature of any error he makes. With practice he may not need to correct his performance for periods of up to 5 seconds or longer, and may even be able to track for a few seconds as accurately with his eyes shut as with his eyes on the display (Poulton, 1957a).

In tracking a single sine wave of very high frequency with a pursuit display, the misalignment error increases as the frequency of the sine wave is increased. With highly practiced men, five different kinds of error have been described (Noble, Fitts, & Warren, 1955): (a) a variable error in phase; (b) a variable error in amplitude; (c) a variable error in the point of termination of successive to-and-fro movements; (d) a constant error in amplitude, which may be an overestimation or underestimation of the track amplitude; the direction presumably depends upon the man's self-instruction, since it can be altered by instruction from the experimenter; and (e) at track frequencies above 3 cycles per second, periods during which the response frequency does not match the track frequency. Again the error may be one of over- or underestimation, and the direction can probably be altered by instruction. The man's difficulty is to tell whether he is ahead or behind, since when, for example, he lags in phase between 180° and 360°, he appears to be leading the track. These results modify and extend those of Ellson and Gray (1948).

In tracking a single sine wave of intermediate frequency the practiced man attempts to match the rate of the track by including an acceleration term in his response, corresponding to the predicted acceleration or deceleration of the track. Errors of misalignment are corrected intermittently. Often the corrections involve changes of rate or acceleration rather than discrete movements altering position. They are easier to detect by special experimental techniques such as intermittent display presentation (Poulton, 1957a).

### *Multiple Sine Waves*

Even a combination of two sine waves can produce a track which appears fairly random, provided the relative frequencies and amplitudes are carefully selected. However, to avoid the possibility of the man learning the sequential properties of the track, it is safer to use three or more sine waves whose phase relationships vary in a nonrepetitive fashion. A more irregular track still is provided by white noise with the higher frequencies filtered out. Some of the learnable characteristics of these quasi-random tracks are the mean position on the display, the mean and approximate maximum amplitude, and the approximate maximum rate of movement. It must be emphasized that no irregular track can be really random if it is to be tracked successfully. Very high frequencies of very large amplitude make tracking virtually impossible, and so have to be excluded.

An irregular track presents a greater problem to the man than a regular track, since it does not appear to repeat. He cannot learn it in detail as he can a simple sine wave. Difficulty can be increased by increasing either the frequency or the amplitude of the track. Doubling either doubles the maximum velocity. But whereas doubling the amplitude practically doubles the size of the tracking error, doubling the frequency may have more or less effect than this, depending upon what the highest frequency in the track is, and upon the nature of the control-system dynamics.

## MULTIPLE TRACKS

Increasing the number of tracking tasks which the man has to carry out simultaneously lowers his standard of performance on them all (Hoffeld, Seidenstein, & Brogden, 1961), because he has to attend to each in turn. When success is scored in terms of the time he is on target on all the tasks, adding an extra task reduces the score both by providing an extra dimension in which he may be off target, and by reducing the time he is on target in all the other dimensions (Bilodeau, 1957).

Tracking in two dimensions with a flat surface as the display and a stylus or joystick control presents a rather special case of multiple tracking, since the two tasks can be integrated into a single task. Instead of having to track first in one dimension and then in the other, the man can track in both dimensions simultaneously. This is only possible when the order of control is the same in both dimensions, so that a single oblique control movement can replace separate movements in the $x$ and $y$ coordinates. Despite the advantage of integrating two tracks into a single display, in principle tracking in two dimensions can never be done quite as accurately as tracking in one dimension. Simply swapping a lever which will move the display marker only in one dimension for a joystick which will move it in two dimensions makes compensatory tracking in one dimension less accurate (Garvey & Taylor, 1959).

The *circular track* of the pursuit rotor presents the simplest tracking task with simultaneous variation in two dimensions. Unfortunately the conventional method of scoring, summed time on target, gives a measure which varies nonlinearly with the size of the target area (Bahrick, Fitts, & Briggs, 1957). Improved scoring methods have entailed compiling the frequency distribution of the durations spent continuously on target (Ammons, 1951), the frequency distribution of the durations spent making noncircular movements (Archer, 1958), and recording continuously the distance of the stylus from the center of the target area (Humphries, 1961), but without specifying the direction of the error. To obtain more complete information an apparatus like Elkind's (1956) "pip-trapper" is required.

## The Display

### PURSUIT AND COMPENSATORY DISPLAYS

The relative advantages of pursuit and compensatory displays in tracking were not investigated until well after World War II (Poulton, 1950b, 1952a). Before then when a choice was available between the two kinds of display, engineering convenience often favored the compensatory (Briggs, 1962a). Yet from the point of view of the man, a *pursuit* display can have four distinct advantages when the track varies continuously with time. Two advantages derive from the fact that the man can see the movements of the track directly, unconfounded by his own response function. (1) He can see the regularities and statistical properties of the track, and thus can learn them. (2) At any time he can predict the immediate future of the track. This is important when the track contains high frequencies, since the man's reaction-time lag then results in a large phase error unless he compensates for it by responding to what he predicts that the track will do next. Prediction is also important if the lag in the control system produces a large phase error (Conklin, 1957). Prediction appears to be based partly upon the visible rate of the track, and partly upon its known regularities and statistical properties (Poulton, 1952b, 1964). It is doubtful whether a man can be sufficiently aware of a transient acceleration to use it in predicting (Gottsdanker, Frick, & Lockard, 1961). More likely he gets his acceleration terms from his knowledge of the regularities and statistical properties of the track. For example, if a sinusoidal track is far from its average position and still moving away from it, there is a high probability that it will be slowing down.

(3) A third advantage of a pursuit display derives from the fact that the man can see his response function directly, unconfounded by track movements. Thus he can learn his control system: what effect a particular control movement or pattern of movements has. And he can learn just what movement or pattern of movements is required to produce a particular effect, since he receives immediate knowledge of results of the effects of his control movements.

(4) The fourth advantage is the status information available in a pursuit display. If an error develops, the man can see whether it has been caused by a shift in the track or in the response function. This can be important in control systems of high order with long lags, such as in the steering systems of ships and in process control in the chemical industry. The many so-called "radar-assisted collisions" between ships in fog (Calvert, 1960) are examples of errors of navigation caused by the lack of status information in the standard compensatory radar display. In the chemical industry it may be important to know whether an error is due to a drift in a control setting which is easily corrected, or to a disturbance in the chemical process itself which may lead to an explosion if it is not stopped.

A *compensatory* display has the four corresponding disadvantages, the seriousness of which can vary according to the nature of the track and of the control system. The disadvantages derive from the fact that the man can see only the error, which is the difference between the track and his response function. As an extreme example, if the control system is complex and of high order, the man may never see the track uncomplicated by some output from the control system. He then has a single equation with two unknown variables. He can learn the control system only if he can produce in the display errors a good deal larger than those produced by the track, so that there is a perceptible correlation between his control movements and the error function. This means that either the experimenter must leave him free to practice without a track until he is reasonably familiar with the control system (Briggs, 1962c, Exp. 3), or else the experimenter must select a track amplitude which is small in proportion to the overall size of the display, and not discourage the man from making large errors until he has learned the control system. For example, the man may "dither," or wobble his control back and forth, to see what effect this has upon the error. Provided he produces higher frequencies than those in the track, he can then see the response of the control system superimposed upon the lower track frequencies.

Once the man has learned the control system, he should be able to learn the track. In the extreme he can produce a steady control-system output, so that any changes in the error are related directly to the track. This is sometimes observed in tracking with a positional control; the man makes sudden discrete corrections separated by intervals of no response during which he can see the track movement. Thus he shares the display time between the track alone and responses which are almost step functions and can therefore be seen when superimposed upon the track movement.

But while the man is actively tracking he cannot use the visible rate of the track to predict its immediate future, as he can with a pursuit display. And he can never be sure whether an error which arises is due to his failure to predict the track sufficiently accurately, or to his failure to make the correct pattern of control movements. Also, unless special training procedures are employed, the man is likely to learn the characteristics of a complex control system and the associated track more slowly than with a pursuit display.

It has been stated as a general rule (Morgan, Cook, Chapanis, & Lund, 1963, p. 220), though apparently without any supporting evidence, that in tracking a compensatory display allows greater magnification than a pursuit display. This is because a pursuit display has to show the full range of possible values of both the track and the control system's response, while a compensatory display need show only a range of values sufficient to encompass all sizes of error which are likely to occur. The implication is that with the compensatory display the sign of the control system's response will always be the opposite to the sign of the track, so that the amplitude of the error is smaller than the amplitude either of the track or of the response of the control system. This is, of course, true of automatic servosystems; but it is not necessarily true in compensatory tracking, since the man may himself generate error which *adds* to the error generated by the track.

In the early stages of compensatory tracking the error produced by the track and by the response function may sum even when the track is predictable and the control system is simple, especially if the direction of control movement required to produce a display movement is unclear (Poulton, 1950b, p. 12). If the control system is complex, the track and the output of the control system are likely to have the same sign from time to time until the control system has been learned, which may take an appreciable period. Until this occurs, the compensatory display needs to cover a larger range than the pursuit, and so has to have the *smaller* magnification. Otherwise the error marker will disappear from view from time to time, and the man may never learn the control system. As we shall see, a nonlinear display magnification gets over this difficulty, but has other disadvantages. The position can be summarized as follows: If the error in tracking with a compensatory display were always as small as the error with the corresponding pursuit display, then the compensatory display could be magnified while the pursuit display cannot be. But in practice the tracking error with a compensatory display can sometimes be very much larger than the error with a pursuit display.

From what has been said, it is difficult to understand how a single compensatory display can ever give better results than a single pursuit display, although it should be possible to produce equally good results by using a predictable track and highly practiced men. Unfortunately most of the experiments in which the two types of display have been compared have used balanced designs, in which the same men have performed a number of different tasks in close succession. Such designs often produce asymmetrical transfer effects between conditions, and are quite unsuitable for this kind of research (Poulton & Freeman, 1965). Where the dynamics of the control system have been held constant over all the conditions performed by any one man, the pursuit display has always proved at least as good as the compensatory, even when the track consisted of an irregular sinusoid with only six or less reversals per minute, and was consequently highly predictable from second to second. This is true both for a positional control system and a rate control system (Heinemann, 1961, Problem C), and also for a simulated aircraft control loop containing rate and acceleration components (Chernikoff, Birmingham, & Taylor, 1956).

When, however, a number of variations in the control system have complicated the experimental design, the compensatory display has twice been found to give better scores than the pursuit display with a low-frequency track. In one experiment positional, rate, and rate-aided control systems were used in random order (Chernikoff & Taylor, 1957). In the other experiment positional, rate, and acceleration control systems were used (Obermayer, Swartz, & Muckler, 1961). These results are difficult to explain on any theory. If switching between control systems made the man less sure of the characteristics of his present system, the pursuit display should have been of more help to him than the compensatory. One can only assume that the confused experimental designs caused the man to adopt a strategy suited to the more difficult compensatory display, and that performance with the normally easier pursuit display suffered accordingly. What form this strategy can have taken is unclear, since the experimenters were not sufficiently aware of the importance of strategies in tracking to consider them worthy of recording.

### QUICKENED AND PREDICTOR DISPLAYS

With complex control systems of relatively high order it is possible to relieve the man of the necessity of learning the dynamics of the system by presenting him with a display which takes them into account. A *quickened* display shows the man how to move his control in order to achieve and maintain the required output of the system (Birmingham & Taylor, 1954, 1961). He has simply to make control movements proportional to the size of the displayed error. The display itself can be pursuit or compensatory, but somewhat different quickening circuits are required in the two cases to take account of the man's tendency to anticipate with the pursuit display and to lag behind with the compensatory. Quickening the display changes the task performed by the man from tracking with a complex control system to tracking with what appears to him to be a positional control system without lag. He can thus greatly reduce the output error of the system.

In one demonstration the man was able to track better simultaneously in four dimensions with quickened displays than in one dimension without quickening (Taylor & Birmingham, 1956). This is because the man has simply to do as he is told; he does not have to understand the system he is controlling in order to control it. He may in fact be tempted not to bother to understand what is happening, which could be disastrous in a sudden emergency. A second disadvantage of a quickened display is shared by automatic systems: inflexibility. The system works well only under conditions similar to those for which the quickening circuits were designed. If the display has been fully quickened, the task does not in fact require a man at all; in flying an airplane the autopilot can be used. This leaves the man free to obtain status information from his displays, to check that the quickening circuits are functioning properly. If he does this, he should be in a better position to deal with a sudden emergency than if he has had to spend his time tracking.

A *predictor* display also relieves the man of the necessity of learning the dynamics of the control system. In tracking in one dimension the ordinate shows the man the difference between the required output of the system and the predicted output, while the abscissa extends this information from the present time into the future. The man can see what future results his present control movements will have, and can adopt control movements which will bring about the required output of the system (Kelley, 1962). This display should reduce the difficulty of tracking with a complex control system, and also reduce the size of the error of the control-system output, but no statistically adequate quantitative results ever appear to have been published. It suffers the same two disadvantages as the quickened display. Some flexibility is available in that the man can select any one of a number of paths to his goal, depending upon how quickly he wishes to achieve it, while with the quickened display only one path is offered. But both display systems can work well only under conditions similar to those for which they were designed.

## DISPLAY MAGNIFICATION

Battig, Nagel, and Brogden (1955) have shown that there is an optimal size of *linear* magnification of a display in compensatory tracking. The experiment used separate groups of men for each size of display, so there is no question of the results being due to a central tendency effect within subjects. When the magnification is too low, the man cannot see the misalignment and rate of movement sufficiently well. This applies to pursuit tracking as well as to compensatory (Hartman & Fitts, 1955). It appears to make little difference whether the decrease in angular size of movement at the eye is achieved by decreasing the size of the movements in the display, or by seating the man further from the display than the optimal distance of 18 inches (Seidenstein, Chernikoff, & Taylor, 1960).

When the magnification is too high, the moving marker is liable to disappear from the face of the display every time the amplitude of the track is too large or the man makes a large error. Thus with a compensatory display and tracks of constant amplitude, higher frequencies which result in larger errors call for smaller magnifications; and early in practice when the errors tend to be large, the amplification has to be less than later in practice when the errors tend to be smaller (Bowen & Chernikoff, 1957).

The marker can be prevented from disappearing from the face of a compensatory display by the use of *nonlinear* display magnification. Here the central part of the display, where the moving marker spends most of its time, is magnified at the expense of the peripheral part where the marker does not go so often. By a skillful selection of the nonlinearity function, it might in principle be possible to obtain the benefit of magnification without the disadvantage of losing the marker. However, the results have so far been disappointing; exponential nonlinear magnification appears to be no improvement on linear magnification (Bowen & Chernikoff, 1957). There are two obvious reasons for this. First, and probably most important, people are not used to nonlinear displays where the center is magnified and the magnification falls off symmetrically in both directions. The control system appears to become "sluggish" whenever the display marker is in the periphery, and "overactive" whenever the display marker is located centrally. Second, the nonlinear function suited to a man in the early stages of practice when his errors have a wide distribution with a long tail will not be optimal for him at a later stage of practice when his error distribution is taller and narrower.

## DISPLAY DEGRADATION

### *Intermittent Displays*

Poulton and Gregory (1952) reported that occluding a pursuit display mechanically for .25 second every 1.25 second produced a reliable deterioration in tracking. The task involved a positional control, and a track consisting of a combination of 3 sinusoids, the fastest of which had a frequency of 25 cycles per minute. No reliable deterioration was found in tracking a single sine wave of 25 cycles per minute, presumably because the track was more predictable. Thus a quasi-random track needs to be seen all the time; occluding the display even for the duration of a blink may lower the standard of performance. This explains why increasing the angular separation of the displays in dual tracking increases the error (Briggs & Howell, 1959); the man then has to look at the two displays alternately.

Illuminating the pursuit display of the Pedestal Sight Manipulation Test intermittently at different frequencies ran into the difficulty that the brightness of the target was affected by the frequency of the illumination (Battig *et al.*, 1954), and that target brightness affected tracking efficiency (Voss, 1955). With brightness held constant, the time on target increased as the rate of flashing increased from 2 to 16 times per second. Since the on-off ratio remained constant at 50:50, the duration of the occlusions fell from .25 to .03 second, at which value the display appeared to be illuminated all the time (Battig, Voss, & Brogden, 1955). Thus where a tracking display is presented on an oscilloscope by the use of time sharing, the frequency of switching should be at least about 16 times per second. Above .25 second the same relationship holds: the longer the duration of the occlusions, the greater the tracking error becomes (Poulton, 1952b; Katz & Spragg, 1955).

If one of the two display markers can be seen all the time, the other only intermittently, the man can do quite well after practice with a positional control as long as it is the track marker which he can see. For his control movements give him an indication of the movements of the response marker. If the track marker is visible only intermittently, seeing the response marker all the time is less help, although it is preferable to seeing both markers only intermittently (Poulton, 1957a).

An intermittent display can also be generated by converting a sinusoidal function into a series of step functions. The larger the steps, the greater the interference with accurate tracking. This holds both in pursuit tracking when the track is shown as a series of step functions (McConnell & Shelly, 1960), and in compensatory tracking when the error is shown as a series of steps (Hunt, 1961).

### *Noisy Displays*

Masking the movements of a display marker by adding visual noise produces a deterioration in the accuracy of both pursuit and compensatory tracking, which increases with the amplitude of the noise. In pursuit tracking with a positional control the effect on the individual display markers is rather similar to the effect of intermittent viewing: noise added to the response marker is least detrimental. If noise is added to the track marker, it makes little difference whether or not noise of similar amplitude is added to the response marker (Howell & Briggs, 1959). In compensatory tracking with a complex display and control system representing the interception of a target by an aircraft, Briggs, Fitts, and Bahrick (1957a) found that noise added to the display did not affect the rate of learning; it merely depressed the level of performance.

#### *Delayed Displays*

Warrick (1949) introduced a simple transmission time lag into the display in compensatory tracking with a positional control. The man could not see his error function until sometime after he had responded. In order to track accurately he had to compensate for the sum of his own reaction time and the display time lag. With a track composed of sine waves of 6 and 30 cycles per minute, a lag as short as .04 second produced a reliable deterioration in performance, although Warrick states that lags of .06 second and less are not normally noticed by the man. However, with a single sine wave of 6 cycles per minute, little deterioration in performance was found with a lag of .08 second. This was presumably because the man had merely to reproduce the sine wave with a constant phase advance.

#### *Filtered Displays*

Garvey, Sweeny, and Birmingham (1958) inserted an exponential time lag into the display in compensatory tracking with a positional control. This not only produced a delay before the man could see the error function, but also attenuated the higher frequencies in the error, whether introduced by the track or by the man's responses. The man thus saw an error function which was both delayed and distorted. The longer the time constant, the less good was the display, and hence the greater was the tracking error.

### Multiple Displays

With two simultaneous tracking tasks requiring separate displays, performance is more accurate when the displays are close together, since head and eye movements are minimized. With two or more displays the fixed zero of the compensatory display is an advantage, since the zeros can be related to each other systematically; tracking is most accurate when for zero error the moving display markers are all aligned and face the same way. With two circular dials arranged one above the other, performance is best with the zeros at 12 o'clock, while with the dials side by side the zeros should be at 9 o'clock (Fitts & Simon, 1952). Even two dials side by side with the zeros fixed at 12 o'clock may prove such an effective combination that compensatory tracking using rate controls becomes more accurate than the corresponding pursuit tracking with wandering zeros (Regan, 1959).

Where two one-dimensional tracking tasks have to be carried out simultaneously, the two displays can be integrated by using the $x$ and $y$ coordinates of a cathode ray oscilloscope. Tracking then becomes more accurate because the two tracks can be seen simultaneously (Briggs & Howell, 1959). If the dynamics of the control systems are the same in the two dimensions, an additional advantage may be gained by reducing the number of display markers and making a single marker move in both dimensions (Chernikoff & LeMay, 1963).

## SENSORY DIMENSION

### *Visual Displays*

The usual tracking display uses the sensory dimension of *visual length* and direction. This is clearly so much better than most other sensory dimensions that few experimental studies have been reported comparing it with others. In some cases experimenters appear to have been satisfied simply by showing that tracking in a particular sensory dimension is feasible.

The best alternative visual display so far discovered is a pair of *flashing lights*, mounted either one above the man's eyes and one below, or one to each side (Brown, Holmqvist, & Woodhouse, 1961). When a light flashed, a joystick had to be moved in the corresponding direction, the size of movement being related to the rate of flash. A light started to flash at a rate of once per second, and was flashing 2.5 times per second by the time the target area had been left. The larger the error, the faster was the rate of flash. In tracking both step functions and an irregular sinusoid, this display was found to be as good as a rather small and poorly designed compensatory display using visual length. Its disadvantage was the usual one of displays using a sensory dimension other than visual length: it was not easy to tell from the rate of flash just how far to move the control. Its advantage was the very clear indication which it gave the man of the direction in which to move his control.

An alternative visual display which is less good under normal conditions involves *brightness*. In compensatory tracking with a positional control, a difference in brightness was found to give a larger mean error than a corresponding difference in length, even after prolonged practice (Moss, 1964a). It was only in peripheral vision that the differential brightness display proved superior. As the angle subtended at the eye between the man's fixation point and the tracking display increased from 0° to 45°, the mean tracking error with both displays increased by a factor of about 4. There was then a small but reliable difference in favor of the differential brightness display (Moss, 1964b). However, a display involving visual length especially designed for use in peripheral vision with large bright markers might have been at least as good.

*Visual movement* is not a good sensory dimension for tracking. Holding a display marker still using a positional or rate control, and moving a response marker at the same rate as a track marker, all gave reliably worse results than tracking with the corresponding display where rate was converted to visual length (Poulton, 1965). Like brightness, movement probably comes into its own in peripheral vision. Movement in peripheral vision is not as good as visual length in central vision (Brown *et al.*, 1961), but it has not been compared with visual length in peripheral vision.

Tracking in the dimension of *visual depth* by means of stereotelevision has been shown to be feasible, but has not been compared directly with conventional tracking (Gould & Smith, 1964).

### *Auditory Displays*

Mowbray and Gebhard (1961) summarized the only comparisons of auditory and visual displays in tracking. The first comparison involved three-state compensatory tracking. With the visual display, being off course to the left or right was signaled by the illumination of a light on the left or right. No light indicated when the subject was on target. With the corresponding auditory display, a sound presented to the *right or left ear* through headphones signaled being off course to right or left. No sound to either ear indicated that the subject was on target. The auditory and visual displays gave equally good results with a positional control after practice, although the auditory display took longer to learn. But when size of error was signaled as well as direction of error, the auditory display fell far behind the visual. This time, with the auditory display "on target" was signaled by a continuous sound in both ears, which may not have been so easy to discriminate from a sound in one ear only. Size of error was indicated by the *rate of interruption* of the sound. The visual display was a cathode-ray tube which showed size of error as a displacement from the target marker. By analogy with the flashing-light display (Brown *et al.*, 1961), performance on the auditory task might have been improved by dividing the target zone into two, and presenting an interrupted sound to the ear corresponding to the position of the error marker in the target zone, instead of the continuous sound to both ears. But whether this would have brought the error down to the level achieved with the display of visual length is problematical.

Feasibility studies only have been made of auditory displays not taking full advantage of the directional information available by using the two ears alternately. All have used constant-velocity or ramp tracks. For *loudness* compensatory tracking involved memorizing a 1000-cycle-per-second tone presented at a loudness level of 40 phons for 3 minutes, and then maintaining this level of loudness using a positional control against changes of .76 decibels per second introduced by the track. This is compensatory tracking without a fixed reference marker. Confining the task to one ear, and presenting a reference tone of 500 cycles per second at a constant loudness level of 40 phons to the other ear, made little difference. The mean tracking error was 2.2 decibels for a maximum track amplitude of 20 decibels. The mean time lag was about 3.0 seconds, which is larger by a factor of well over 10 than the corresponding time lag in tracking a compensatory display of visual length. Pursuit tracking involved adjusting the intensity of a 500 cycle-per-second tone to match the varying loudness of a 1,000 cycle-per-second tone. It made little difference whether the two tones were presented together, or to separate ears so that the combined stimulus could be centered in the head. Here the mean tracking error was 3.3 decibels (Pikler & Harris, 1960).

For *pitch* compensatory tracking involved memorizing a tone of 250 cycles per second presented for a minute, and then compensating for a track containing ramps intermittently of 1.25, 2.5, or 5 cps per second. Even after practice tracking was so erratic that only two measures were taken: (a) The time lag at a change of rate in the track. This was largely determined by the difference of 5 cycles per second which is required at this frequency to discriminate an absolute difference in pitch. Thus the time lag was longer with the slower tracks. (b) The constant error in pitch when the track and the response function were both stationary. This was about 3 cycles per second after 5 minutes of tracking. It was independent of the rate of the track, but increased with time after memorization of the 250 cycle-per-second tone. Presenting a constant musical E to one ear did not help in trying to maintain a constant musical C in the other ear against disturbances introduced by the track. Pursuit tracking, in which a musical C presented to one ear had to be kept in step with changes in the musical E presented to the other, was quite impossible for some people. Apparently it generally takes several seconds to set up the musical interval, and this could not be done while the musical E was changing pitch (Harris, Pikler, & Murphy, 1963).

Compensatory tracking in a simulated aircraft has also been carried out with an auditory signal varying simultaneously in *three sensory dimensions:* apparent movement from one side to the other, a rise or fall in pitch, and a rate of interruption. After training it was claimed that people could track as well with this combined auditory display as by obtaining information separately about the three dimensions from a set of visual displays (Forbes, 1946). However, a properly integrated visual display, which would have been more comparable to the combined auditory display, was not used.

*Vibratory Displays*

A compensatory vibratory display using apparent movement over the skin has been constructed by strapping three vibrators to the chest wall. The direction in which to move the control was indicated by the order of activation of the vibrators, whether from left to right or from right to left. The larger the error, the more frequently the composite signal occurred. Tracking with this display was found to be at least as good as tracking with the corresponding visual display of 3 lights (Geldard, 1961); but as we have seen, visual movement is not a good sensory dimension for tracking.

Vibrators have been attached to the thumb and index finger holding a lever control, to indicate the direction only of the rate of change of the error. The man had to track in one dimension with a complex control system representing an aircraft. The track consisted of step functions, and he had a compensatory visual display as well as the vibrators. Tracking with the vibrators was found to be more accurate than tracking without (Hirsch & Kadushin, 1964). Apparently the value of the vibrators was in calling attention to small changes in rate when the visual display was almost stationary. With a control system of high order, this reduced the size and number of the oscillations about the target.

## The Control

### KIND OF CONTROL

Most precise tracking is achieved using the muscles of the hand and wrist, as in writing. When the task is easy there may not be much to choose between a *joystick* operated by the thumb, the hand, or the forearm. But under difficult conditions of a sensitive rate control with an exponential time lag of 2.0 seconds, the hand joystick is found to be reliably the quickest in step-function tracking, and the forearm the slowest (Hammerton & Tickner, 1966). This is presumably because the hand is designed and used all the time for precise movements. The thumb, which is used in gripping, does not normally have to make such precise movements. And the forearm is designed as much for strength as for precision. Foot controls permit the use of still greater force, but with still less precision (Morgan *et al.*, 1963, p. 265).

For pursuit step-function tracking in two dimensions with a positional control system without lag, a *stylus* control held like a pencil and actually moved over the surface of a cathode-ray tube display is about twice as quick as a 6-inch joystick operated by the hand (Baker, 1960, p. 157). The two advantages of the stylus are, first, that it uses the greatly overlearned movements of writing and drawing and, second, the tip of the stylus either is itself the display marker, or else it is located over the controlled display marker. This is the most compatible control-display relationship other than moving the display marker with the fingers (see below). Where the movement of the display marker does not correspond directly to the movement of the stylus, the two cannot remain superimposed and the advantage of almost maximal compatibility is lost. This is the case in compensatory tracking, and in pursuit tracking where the dynamics of the control system transform the direct output of the control. It might then be equally good to let the stylus write on a separate cathode-ray tube, as Elkind did (1956), although nobody appears to have checked this.

A *rolling-ball* control of 5-inch diameter is not much slower than a 6-inch joystick operated by the hand for control displacements of up to 180° of ball rotation. But it becomes progressively worse for larger displacements (Thornton, 1954).

## COMPATIBILITY OF CONTROL WITH DISPLAY

The perfectly compatible control is the display marker itself. When the man wants to move the marker, he simply takes hold of it and moves it directly. The nearer the control can approach to this ideal, the simpler the coding problem becomes for the man. In order to minimize the coding problem the control should be mounted as close as possible to the display, should move in the same direction as the display marker, and should have the minimum of control-system dynamics upsetting the direct relationship between control and display movement.

Compatibility is particularly important when the display is compensatory, since the only way in which the man can diagnose the nature of the disturbances fed into the error function by the track is by computing the control-system output function and subtracting it from the displayed error function. This may be too difficult if a complex transformation is required in order to compute the output function from the control movements. With a control system of high order, partially quickening the display by adding a positional component produces a directional compatibility between control and display movement which can improve performance (Garvey & Taylor, 1959, Exp. 1). A disadvantage of tracking with a rate control system, or with a control system having rate as its lowest component, is that the directional relationship between the movement of the control and the output of the control system is likely to reverse twice per cycle. The same holds for a minus-first-order or differentiating control system (Poulton, 1963a).

A single joystick is more compatible with a two-dimensional display, as on a cathode-ray oscilloscope, than are two levers for movements in the $x$ and $y$ planes, respectively. Thus where the dynamics of the control system are the same in both dimensions, tracking is more accurate with the single joystick. But where the dynamics are different for the two dimensions and do not interact, the man has in effect two separate tasks to perform. In this case separate controls are best (Chernikoff & LeMay, 1963), presumably because they help the man to distinguish between the two control systems.

## CONTROL GAIN

The classical work on control gain was carried out by Jenkins and Connor (1949) using a rotary positional control and step functions. They found that there was an optimal amount of gain for a display movement of a given size. If the control had to be rotated much further than this, it took too long to bring the display marker up close to the target. If the control had to be rotated a good deal less than this, the final alignment took too long, since a relatively small movement of the control produced an unduly large movement of the display marker. The best value of gain produced a relatively quick movement of the display marker to the region of the target, followed by a relatively quick final alignment. Changes of gain in either direction led to slower acquisition of the target, so that the relationship between acquisition time and control gain was U shaped.

The conceptual model of an optimal control gain has to be modified when a joystick control is used, since a joystick does not normally move further than 90° on either side of its central position. It is almost as quick to move a joystick through 90° as through a smaller distance. There is nothing to correspond to the numerous full rotations of the control knob which were sometimes required in the classical experiment. Thus with a joystick positional control the optimal gain is obtained by equating the maximal joystick movement with the maximal required movement of the display marker, and the relationship between acquisition time and control gain is J shaped (Jenkins & Olson, 1952). For the same reason, with control systems of higher order the maximal joystick movement should be equated with the maximal system output (Hammerton, 1962).

## CONTROL RESISTANCE

Howland and Noble (1953) compared different combinations of control resistance in compensatory tracking of a 15 cycle-per-minute sine wave. They found that simple spring centering gave the greatest time on target. Next best was spring centering combined with viscous resistance, and then viscous resistance alone. All these conditions were better than a control with minimal resistance. Inertial resistance gave the smallest time on target, whether by itself or combined with the other forms of resistance.

The advantage of spring centering is that the spring resistance gives the man a sensory indication or "feel" of the amplitude of his control movements. It is particularly useful to him where he cannot obtain the information directly from the display. This is the case when he has a compensatory display, as in the experiment of Howland and Noble, when he has to alternate his attention between two tracking tasks (Briggs & Wiener, 1959), and when complex control-system dynamics intervene between a control movement and the corresponding movement in the display. In compensatory tracking with complex dynamics representing an aircraft, Briggs *et al.* (1957b) found that a spring-centered control which had an appreciable amplitude of movement was reliably better than a spring-centered "pressure" control which did not. Thus the easiest all-round control for the man to use is one in which both its position and its spring resistance indicate to him his present input into the control system.

### *Control Resistance for Special Tracks*

In a theoretical paper Bahrick (1957) stated that spring resistance is most suited to accurate positioning, viscous resistance is most suited to achieving an accurate rate of output, and inertial resistance is most suited to achieving an accurate amount of acceleration. This is because the man uses force as a cue, and a given force produces a given displacement with spring resistance, a given rate of displacement with viscous resistance, and a given acceleration with inertial resistance. The usefulness of viscous and inertial resistance was demonstrated in an experiment in which the man first made circular and triangular movements with a visual guide and a metronome, and then reproduced them without. Both kinds of control resistance produced greater uniformity in the speed of movement and in the peak velocity (Bahrick, Fitts, & Schneider, 1955). However, more practice was required before the value of spring resistance in producing greater accuracy of positioning could be demonstrated (Bahrick, Bennett, & Fitts, 1955). A theoretical difficulty in interpreting the results is that the arm itself supplies both viscous and inertial resistance, and this has to be added to that of the control.

If Bahrick is correct, two combinations of control resistance and control-system dynamics are suitable for constant-velocity tracks: a spring-centered rate control, and a positional control with viscous resistance. And three combinations are suitable for constant-acceleration tracks: a spring-centered acceleration control, a rate control with viscous resistance, and a positional control with inertial resistance. It is not known which combination is in fact the easiest to handle.

### *Other Uses of Control Resistance*

Friction, a constant resistance which has to be overcome before the control can be moved, generally lowers the standard of tracking (Helson, 1949). Unlike spring centering, friction gives the man no indication of the position of his control. And when he reverses the direction of his control movement, there is a kind of dead space corresponding to twice the force required to overcome the friction, over which changes in force produce no change in the position of the control.

However, in designing a control for a particular purpose, control resistance may be required for other reasons. Control resistance helps to prevent accidental operation of the control, or else helps to minimize the effects of accidental operation. Also control resistance helps to filter out tremor and sudden jerks in the controlling limb, which may result from vibration or sudden jolting of the vehicle in which the control is mounted (Morgan *et al.*, 1963, p. 256).

## The Control System

### ORDER OF CONTROL

The control system most suited to human capabilities is zero order or *positional*, for there is an immediate and obvious relationship between control and display movements. In order to acquire a target it is only necessary to move the control the corresponding distance in the correct direction, whereas with a first-order or *rate-control* system two movements are necessary: one to put on the rate of movement, and another to take it off again when the target has been reached. With a second-order or *acceleration-control* system at least three movements are required: one to put on the acceleration which in time produces the rate of movement, one to take off the rate of movement again, and then a final movement to return the control to the central position. Also movements after the first have to be correctly timed. It is necessary to start to remove the acceleration a little before the target is reached, or the rate of the movement of the display indicator will cause it to overshoot. And the control has to be returned to the central position as the indicator comes to rest, or it will be left with a residual velocity. With a third-order or "*jerk*"-*control* system, a still more complex series of control movements is necessary to acquire the target. With a minus-first-order or *differentiating-control* system the target can be acquired simply by moving the control at the correct rate; but the rate has to be maintained or the display indicator returns to its starting point.

An *aided* control system combines two or more orders of control. For example, a control movement may move the display indicator directly, and also give it a rate of movement in the same direction. If a control system has to have components of higher order, it should also have a positional component if at all possible; thus it should be aided rather than pure rate. A pure rate-control system is suited only to tracks of practically constant velocity. For a sinusoidal track to give better results with a rate control, the highest frequency in the track has to be so low that the sinusoidal track almost resembles a constant-velocity track.

The much-quoted experiment (Chernikoff & Taylor, 1957), from which it is claimed that a rate-control system is easier than a positional system for frequencies below about 10 cycles per minute, should never have been designed with all experimental subjects using the positional, aided, and rate controls in balanced orders. For Lincoln (1953) had already shown that, whereas there is positive transfer in changing from a positional-control system to either of the other two, and when changing between aided- and rate-control systems, there can be *negative* transfer in changing from either of these two to the positional-control system. At a guess, in the absence of transfer effects the highest frequency in the track has to be 1 to 2 cycles per minute or less before a positional-control system ceases to be optimal.

With track frequencies of 40 cycles per minute, a minus-first-order control system which differentiates the man's response is almost as good as a zero-order or positional-control system (Poulton, 1963a,b). This is because with a differentiating control system the man can change the position of the display marker simply by changing the rate of his control movement.

## CONTROL TIME LAGS

### *Transmission Lag*

With a simple transmission time lag the control system does not respond when the man first moves his control, and then the normal effect of the control movement appears. A transmission lag limits the frequencies of an irregular track which can be dealt with. If a quarter of a cycle elapses during the time lag, the track is correcting itself at this frequency before the corresponding control movement can take effect. If a full half-cycle elapses during the time lag, a control movement aimed at correcting a track displacement at this frequency will in all probability produce an error of the opposite sign. The only way in which the man can compensate for the time lag is by knowing what the track will be doing at the corresponding time ahead, and responding accordingly. This is only possible with predictable tracks.

### *Exponential lag*

This is the commonest kind of time lag in electronic equipment. Something happens almost as soon as the man moves his control, but the full effect of the control movement only appears gradually, following an exponential time function. In tracking step functions with a lagged positional control system, the man has to move his control further than he does when using a control system without a lag, in order to reach the target reasonably quickly. An exponential time lag is thus somewhat similar to a reduction in control gain. When these two independent variables are varied simultaneously on the same experimental subjects, the optimal gain is found to be higher with long lags than with short (Rockway, 1954).

Having produced a relatively quick movement of the display indicator by a large control movement, the man then has to avoid the overshooting which will gradually occur as a result of the remainder of the exponential time function. This means moving the control back to the other side of its final position as determined by the size of the step, in order to cancel the remainder of the initial exponential time function; and then gradually moving the control out again to its final position. The three successive movements, far out, back a good way, and then finally out a bit further, correspond in direction to the movements required in step-function tracking with an acceleration control.

### *Sigmoid Lag*

This behaves rather like a combination of a transmission lag and an exponential lag. Nothing much happens when a control movement is first made, and then the effect appears gradually.

The classical work on exponential and sigmoid time lags was carried out by Conklin (1957), who used a positional control with time constants of up to 16 seconds. The longer the lag, the worse the man did, but performance was improved by any condition which assisted prediction. Thus performance was better with a single sine wave than with a combination of two or three sine waves; and all these tracks gave better results than a track consisting of low frequency white noise. Performance was also better with a pursuit display than with a compensatory display. And the exponential time lag, which gave an almost immediate indication of a control movement, was preferable to the sigmoid time lag which did not.

### *Time Lag and Order of Control*

The order of control of a ship or chemical plant is determined by its physical design. However, if suitable sensing devices are available, it may be possible to lower the order of control, inserting a time lag instead. For example, in steering a ship down a canal, the rudder acts like an acceleration control; it makes the ship turn at a corresponding rate, and the turn has to be maintained for a while before the ship appreciably alters its distance from the side. If the angle which the ship makes with the side of the canal can be sensed directly, it is possible to build the helmsman a lagged rate-control system. With this the man simply selects the required angle, and after the time needed to move the rudder and to turn the ship, the ship assumes the corresponding angle. When the ship has moved most of the required distance toward the side of the canal, the man centers his control again, and the ship gradually turns back onto its course parallel with the side of the canal.

The lagged rate-control system simplifies the man's job only if he is not in a hurry. If he has to change the distance of the ship from the side of the canal as quickly as possible, and the control system allows this to be done, he has to move his control sufficiently far to turn the ship in time through 90° or even 180°. Long before it turns as far as this the ship will have moved most of the required distance toward the side of the canal. To prevent the ship from going too far he has then to move his control well over to the other side of its central position, before finally returning it to the center. This is a similar sequence of movements to that required with an acceleration control.

If the distance of the ship from the side of the canal can also be sensed directly, it is possible to build the helmsman a still more highly lagged positional-control system. With this the man simply selects his distance from the side of the canal. The lag then includes the time needed to move the rudder, to turn the ship, to take the ship closer to the side, to turn the ship back onto its course down the canal, and to return the rudder to its central position. This automatic system also only simplifies the helmsman's job if he can accept the long delay. If he is in a hurry, and the control system allows it, he still has to make three control movements: first selecting a position too near to the edge of the canal, then a position less far over than the required end point, and finally returning his control to the position corresponding to the end point.

Process control in a chemical or petrochemical plant can combine a high order of control with a long time lag. This presents the man with a particularly difficult task, worse even than controlling a slow massive ship (Goldstein & Newton, 1962). By the time the man can see that a control movement is taking effect, it may already be too late to prevent the process from overshooting the required target even if he puts his control hard over in the opposite direction. It is therefore necessary to operate the control without any direct feedback from the plant; it may be 15 or even 30 minutes before the man can find out whether or not his control movements have achieved their purpose. If the task cannot be automated, the man may be assisted by a quickened or predictor display (see above).

## Control System Output and Control Gain

In most engineering control systems the maximum output is fixed by the design. For example, a car has only a certain power of acceleration. The only method of increasing the control gain of the accelerator pedal is to let full power be obtained with the pedal only partly depressed, instead of fully depressed. As we have seen, this may very slightly reduce the time taken by the foot to move the pedal from its zero position to the position of full power. Its main effect is to reduce the foot's ability to select just the required amount of power. Thus in tracking the optimal control gain is achieved by equating full power with the full range of movement of the control, and the relationship between acquisition time and control gain is J shaped (Hammerton, 1962).

In contrast, in the laboratory the gain of a control is conventionally increased by simultaneously increasing the output of the control system. To return to the example of the car, increased power is made available on full depression of the foot pedal. In addition to a more sensitive control, the man thus has more power to use if he needs it. Any difficulty caused by an oversensitive control may thus be more than offset by the additional power. This produces a U-shaped relationship between acquisition time and system output, even with a joystick control. Too little power prevents the display marker from moving fast enough, while too much power hinders fine acquisition movements.

## CONTROL SYSTEM DEGRADATION

### *Intermittent Control Systems*

In contrast to the conventional analog control system, a digital computer can accept the man's output only at finite time intervals. Bennett (1956) investigated rates of acceptance of the output of a velocity control ranging from twice to 8 times per second. Each time the output of the control was accepted, the marker of the compensatory display altered its rate of movement according to the distance moved by the control during the interval since the last acceptance. With both step and constant-rate tracks, longer intervals gave worse performance. An added transmission lag was particularly detrimental with longer intervals between acceptances.

### *Oscillatory Control Systems*

Owing to their physical design, high-performance aircraft, and space craft during re-entry into the earth's atmosphere, tend to oscillate as the pilot attempts to take them from one steady state to another. The long-period or phugoid oscillations are not very highly damped, and are difficult to control since movements of the control stick aimed at tracking them out produce further oscillations which in turn call for further control movements. With a compensatory display and a track consisting of a sine wave of 6 cycles per minute, Muckler, Obermayer, Hanlon, and Serio (1961b, c) found that the largest period of phugoid oscillations with which tracking performance was above the no-response level depended upon the size of the control movements required. With a low-amplitude track and high control gain, students could deal with a period of oscillation of 71 seconds. But when the amplitude of the track was doubled and the control gain was halved, they could deal adequately with a period of oscillation of only 18 seconds.

## The Man

### Craik's Theory of Intermittency

Craik (1947, 1948) put forward the suggestion that in tracking a man behaves like an intermittent ballistic servo. He responds only intermittently, and his responses are made without the aid of voluntary monitoring until they are complete. The rationale underlying the theory has been dealt with at greater length in the preceding pages: it takes a minimum of about .20 second for a man to respond to a visual stimulus. The minimum for a kinesthetic stimulus is somewhat shorter, about .15 second (Chernikoff & Taylor, 1952), but people are not normally set to respond to kinesthetic stimuli. During the early part of the reaction time it may on occasions be possible to inhibit the response partly or completely (Vince, 1948a, Exp. 4). But later during the reaction time the computing system in the brain becomes refractory. It then takes a mean of about .30 second before the response can be amended (Hick, 1949). The time is made up of the minimum of approximately .10 second required, as it were, to clear the computing mechanism before it can accept the next input (Poulton, 1950a, Fig. 4; Davis, 1956), and the .20 second visual reaction time required to process the input.

Craik's theory follows from this. First, there is a minimum interval with a mean of about .30 second between responses in tracking. For compensatory tracking in the early stages of practice with a constant-rate track, the interval averages about .50 second, most times lying between .25 and 1.00 second (Vince, 1948a, Fig. 2). The interval does not change appreciably with changes in the magnification of the display. Nor does it alter with changes in the rate of the track, the mean tracking error being proportional to the rate of the track (Vince, 1948a). The difference between the minimum interval between responses and the observed interval is accounted for by assuming that the man normally likes to see the effect of his last response before initiating his next one.

Second, since it requires about .30 second before a response can be amended (Hick, 1949), quick responses which take .30 second or less to complete have to be made ballistically, or without the aid of visual monitoring. They are no more accurate when the man can see his movement than they are when his vision is obscured momentarily (Vince, 1948b, Fig. 4).

Crucial to Craik's theory is the concept of the point of no return during the reaction time. Once this point has been passed the man cannot prevent his response from being made, and a fresh reaction time is required before he can do anything else. The lengthening of the subsequent reaction time, due to so-called "psychological refractoriness," is not crucial to the theory, although it helps to account for the observed time relationships (Welford, 1952, 1959).

A point of no return is essential from an engineering point of view (Craik, 1947). A servomechanism with a transmission lag of .15 to .20 second would oscillate all the time if the amplitude of the movement it produced, the rate, or the acceleration were determined continuously by the misalignment. For when it had reached the point of zero misalignment it would still be responding to the misalignment .15 to .20 second previously, and so would overshoot. This tendency could be overcome if a misalignment triggered off a ballistic movement accurate to, say, about 10 percent (Vince, 1948b, Fig. 4), and the next movement was not triggered off until the previous one had been completed.

### *Strategies to Deal with Intermittency*

In his studies of compensatory tracking with a positional control and constant-velocity track using unpracticed people, Craik (1947, p. 59) noted that the intermittent corrections of misalignment soon came to have superimposed upon them a constant rate of movement. This corresponds to the "rate aiding" of the servo engineer, and suggests that the man was aware that the track had a constant velocity. He, as it were, adopted a program which seemed to him to match the track, and then simply corrected for the error between the program and the track. The rate element in the man's response could be demonstrated by unexpectedly stopping the track. The man continued to respond for a reaction time before he realized that his program was no longer appropriate (Craik, 1947, p. 58; Vince, 1948a, Fig. 1C).

In tracking with a positional control, the intermittent corrections tend to blend with the program and become obscured with practice (Poulton, 1957a, p. 92). They are most easily distinguished by using a track which in no way resembles a series of discrete corrections. For this purpose the constant-velocity track used by Craik and by Ellson, Hill, and Gray (1947), and the relatively fast simple-harmonic track used by Noble, Fitts, and Warren (1955) and by Poulton (1957a) are preferable to a complex harmonic track. It is then found that after practice the man may not correct his program as often as twice per second. He may wait until an appreciable error has developed, which can take several seconds (Ellson, *et al.*, 1947; Poulton, 1957a).

In paced contour tracking also the man may not respond as frequently as twice per second. When he can see the wiggly line sufficiently far ahead, he tends to aim first at one bend and then at the next. If he has time, he may correct himself when he badly overshoots or undershoots at a bend. But he tends not to worry too much about his error between the bends (Poulton, 1964).

## STRATEGIES TO DEAL WITH REACTION-TIME LAG

In tracking the man always carries the handicap of his reaction-time delay. With tracks of low frequency and moderate amplitude, this handicap of about .2 second does not matter much. For the track will not have moved very far during this time. The resulting error is therefore likely to be small compared with the error introduced by the man from other causes. The handicap is also not serious with control systems having long time lags, since .2 second will be only a small percentage of the total lag. In these cases the man does not need to adopt a special strategy to reduce his own time lag.

But with tracks of high frequency and control systems without appreciable time lag, a human transmission lag of .2 second can be very damaging to performance. For although the man may be able to reproduce the track reasonably accurately after a delay of .2 second, his response function will be out of phase with the track. At a track frequency of 75 cycles per minute he will be 90° out of phase. This means that on the average half the time he will be correcting a displacement of the track which is already correcting itself, and the other half of the time he will be increasing the size of the error caused by a track displacement. Some strategy is therefore necessary to reduce his reaction-time lag.

Adopting an extreme strategy, the man can aim to time his responses correctly, regardless of their accuracy. In contour tracking where he can see the track ahead, he has simply to reproduce the track as he sees it one reaction time ahead of his response indicator, which he can do reasonably accurately. But in pursuit and compensatory tracking where the track is not visible ahead, he has to predict the track one reaction time ahead, and respond according to this prediction. With regular tracks such as simple sine waves and simple ramp tracks with constant slopes, he can predict reasonably accurately with practice. Thus he can combine accuracy in both the time and displacement dimensions.

But with irregular tracks accurate predictions cannot be made .2 second ahead, although the man can often predict better than chance after practice as long as the frequencies in the track are not too high. In these circumstances he tends to adopt a compromise strategy. He accepts some degree of error in prediction in order to reduce his time lag (Poulton, 1962). The higher the frequencies in the track, the more damaging his time lag becomes, and hence the more extreme the strategy he is likely to adopt to minimize it.

Very high-frequency irregular tracks cannot be dealt with adequately at all. For the man's ability to predict the track any appreciable time ahead is not far above the chance level. And if he copies the track as he sees it, his reaction-time lag means that he will be just as likely to increase the error produced by the displacement of the track as to reduce it. The ideal electronic analogy to the man's behavior is thus a low-pass filter since he can transmit low frequencies but not very high frequencies. However, this analogy must be used only with caution, since the presence of high frequencies in the track tends to mask the lower frequencies, and so reduces the man's ability to track the lower frequencies (Howell & Briggs, 1959).

## Spare Channel Capacity

In tracking in one dimension the typical display consists simply of two markers, one or both of which move over and back. Right at the start the task appears to the man to be interesting and challenging. But soon he discovers what he can and cannot achieve, and settles down to give what he considers to be an adequate performance. A small error comes to be tolerated, and whatever effort is made is directed at preventing or correcting large errors (Helson, 1949, p. 495). The task is analogous to a vigilance task, and fails to occupy the man's full channel capacity or attention. In the case of the pursuit rotor, the suboptimal performance has been attributed to conditioned inhibition (Kimble, 1949).

At this stage, the level of performance can be improved by presenting the man with a challenge. For example, with a compensatory display the size of the error can be magnified. To an experienced man this appears to indicate poorer performance, and he tracks more accurately as a result (Helson, 1949). Alternatively, the man can be presented automatically with an auditory signal whenever the error is below a critical value (Reynolds & Adams, 1953a); or he can be given summary feedback after each trial (Karlin & Mortimer, 1963). Both forms of knowledge of results can reduce the size of the error which the man will tolerate, and so raise the standard of performance.

Unfortunately, a change in experimental conditions which makes the task harder may also present a challenge to the man. This means that the lower level of performance which is to be expected as a result of the increased difficulty of the task may be partly offset by the challenge effect. Tracking in one dimension is thus not as sensitive to changes in experimental conditions as are tasks which occupy the man's channel capacity more fully, for example, tracking in two dimensions (Frost, 1962).

Sensitivity can be improved by reducing the man's reserve channel capacity. A number of different so-called "task-induced" stresses will do this. Garvey and Taylor (1959, Exp. 3) compared a positional-control system with an acceleration-control system using a compensatory display and a track consisting of sine waves of 3.2 and 2.3 cycles per minute. After practice performance was equally good with both control systems, although the acceleration control is clearly the more difficult system to use. The difference in difficulty was revealed by (1) leaving the man to track for an hour nonstop; (2) reversing the usual directional relationship of control and display movement; (3) making the man use his nonpreferred hand; (4) rotating the plane of track movement in the display through 45°, and giving the man a joystick free to control the display in two dimensions instead of a lever moving only in one dimension; (5) giving the man an identical tracking task to perform with his other hand; (6) giving him an additional visual task to perform simultaneously; (7) giving him an additional auditory task to perform simultaneously.

In addition to acting as a challenge to the man with the easy control system, the task-induced stress appears to remove the effect of practice with the difficult control system, and thus reveals the difference in difficulty between the tasks (Garvey, 1960; Fuchs, 1962). At the start of practice in compensatory tracking the man responds principally to the size of the error which he can see, and gives little weight to the rate of change of the error, whereas after practice he takes the rate of change of the error more into account, and even gives some weight to the acceleration or deceleration of the error. At this stage the man is behaving somewhat like an acceleration-aided servosystem, and the computations required to do so occupy more channel capacity than is required to behave as he does initially like a simpler servosystem. The task-induced stress removes the additional channel capacity, and the man returns to his original behavior of responding principally to the error which he can see. In the case of the Garvey and Taylor (1959) experiment this simple behavior is more appropriate to a positional-control system than to an acceleration-control system. When the stress is applied, performance with the positional control system thus suffers considerably less than performance with the acceleration control system, and it also benefits more from the challenge effect.

Both a task-induced stress and magnifying the size of the error in a compensatory display challenge the man. Since there is an upper limit to what the man is capable of doing, the combination of a task-induced stress and error magnification should have less effect than the sum of the two separate effects. This has been found to be the case (Garvey & Henson, 1959).

## HUMAN TRANSFER FUNCTIONS

Engineers have developed mathematical techniques for analyzing the behavior of a nonlinear closed-loop servomechanism (Graham & McRuer, 1961), and these have been used to analyze human tracking behavior (McRuer & Krendel, 1959). Like the nonlinear servomechanism, the man does not respond exclusively in ways which correlate with the track. The human transfer function thus has to have two parts: a *describing function* which shows the best linear fit between the response and the track, and a *remnant* term which includes those aspects of the man's performance which are not correlated with the track.

In using this approach it is desirable to minimize the size of the remnant term, so that as much of the behavior as possible is fitted by the describing function. This is done by selecting values of the engineering variables which make the man's response as nearly as possible a linear fit of the track, and thus as close as possible to the behavior of a linear servomechanism. Changes of strategy resulting from the man's ability to perceive the track and relate it to previously stored information increase the size of the remnant and are excluded if possible. The man is usually given only compensatory displays, in which the track is disguised by being combined with the response function. If the dynamics of the control system are sufficiently complex, the disguise may be adequate by itself even when the track is a single sine wave. But with less complex dynamics it is also necessary that the track appear to the man to be random.

Clearly the more successful the describing function is in fitting the man's behavior, the closer the behavior must be to that of a linear servomechanism, and the more that is characteristically human about the behavior which has had to be excluded from the study. Thus by its very nature a describing function can never be more than a partial description of the full range of human tracking behavior.

### *The Describing Function*

The most complete describing function for the human being contains six terms (McRuer & Krendel, 1959):

(1) A *gain* term, relating the amplitudes of the track and response function. It is likely to vary according to the instructions given to the man, or according to what he thinks he is supposed to do if he is not specifically instructed in this respect. Changes in the engineering variables which make the task so difficult that the man can see that he is inserting more error than he is tracking out tend to reduce his gain. Examples are an increase in track frequency, an increase in control-system time lag, and an added control-system integration. When the gain of the control system is changed, the man eventually compensates for this so that the combined gain of the man and control system remains approximately constant.

(2) A *reaction-time* term. The human time lag may be almost zero if the man realizes that he is faced with a highly predictable track such as a pro-

longed ramp function or a single sine wave. But with a random-appearing track the average value is about .15 second. It is very much a function of the individual man, ranging from about .12 to .20 second.

(3) A *neuromuscular lag* term, due to the inertia of the responding limb. The lag itself can presumably range from some small value when a pressure control without appreciable movement is operated by the hand or thumb, to a value of perhaps .1 second when a whole limb has to be moved a good distance.

(4) A *first-order lead* term, said to act like anticipation, and to contribute particularly to stability at high frequencies. The lead term appears when an appreciable lag is introduced by the control system, and when the control system is changed from zero order to first order by the insertion of an integrator. The term is presumably related to the compromise strategy which the man adopts in dealing with time lags. With tracks of high frequency a time lag anywhere approaching the man's reaction-time lag of .2 second is likely to have a disastrous effect upon the error, since the response may be 90° or even 180° out of phase with the track. The same is true with tracks of intermediate frequency if the control system introduces a large time lag. The man therefore tends to respond to what he predicts that the track will be doing next. In this way he reduces the time lag, and hence the phase error, but the response function becomes a less accurate copy of the track. The higher the frequencies in the track, the greater will be the tendency to reduce the time lag at the cost of a reduction in accuracy.

(5) A *first-order lag* term, said to contribute particularly to stability at low frequencies. With tracks of low frequency a reaction-time lag of .2 second may hardly be noticed by the man, since it produces such a small phase error. He may lag still further behind the track in order to reproduce it with greater precision.

(6) An *indifference threshold*, due to the fact that the man does not always attempt to correct small errors in either direction. He behaves as if he had a conceptual "target area," and only attempts to correct excursions outside this area, or movements which appear to him to be leading to excursions outside the area. Normally the indifference threshold has only a minor effect upon the describing function.

Once the numerical values of a describing function for the human being have been fitted, they appear to hold within the limits of experimental error over a range of amplitudes of track, and over a range of display magnifications. But any major change in the engineering variables calls for new values. Thus the describing function for a high-frequency track will not have the same values as that for a track containing only low frequencies. Also, changing the display from pursuit to compensatory, changing the type of control, or changing the parameters of the control system, all call for new fitted numerical values. Changes in the man also mean new fitted values. Any set of numerical values will hold only for one particular man, at one particular stage of learning, with one particular degree of fatigue or reactive inhibition, and when he is using one particular strategy. A change in any of these variables means a rather different set of numerical values in the describing function.

This makes good sense psychologically, since changes in any of the variables listed above would be expected to result in changes in performance, which is what the describing function reflects. But it means that although the general form of the describing function for the human being may be known, the numerical values have to be calculated for each detailed application. The six terms in the describing function in a sense correspond to the intervening variables of the learning theorist. Clearly one of the next steps for engineers wishing to proceed further with describing functions for man is to discover by experiment the range of values which each of the six terms can assume, and show in greater detail how these values are affected by changes in the relevant independent variables.

### *The Remnant*

This represents that part of the response in tracking which is not related linearly to the track. Models accounting for the remnant fall into one of three general classes (McRuer & Krendel, 1959):

(1) *Nonlinear transfer.* Like the transfer characteristic of the nonlinear servomechanism, the human transfer characteristic may be nonlinear. If, as seems likely, the human response in tracking is basically intermittent, it is bound to contain nonlinearities when treated as a continuous function. Against this as a complete explanation is the finding that the size of the remnant varies systematically with changes in the engineering variables, for example, with the complexity of the control system dynamics, and with the amplitude and frequency of the track. At least some of these variations can be predicted from the other models.

(2) *Nonsteady operator behavior.* This may occur during the sample tracking period of 2–4 minutes required to calculate the transfer function, despite the experimenter's attempts to prevent it. The man may become fatigued; many experimenters deliberately restrict periods of tracking to 45 or 60 seconds to avoid noticeable fatigue effects. Or the man may think that he can anticipate the track and respond accordingly, instead of responding to the displayed error. "Dither" is another example of a response not called for by the changes in the error. Here the man wobbles his control back and forth while he is responding, so that his lower-frequency responses to the track have a higher-frequency oscillation superimposed upon them. If the

experimenter notices that the man is doing this, he can get rid of the behavior by requesting him to stop it.

Nonsteady behavior which cannot be prevented may occur in two-dimensional tracking. Here the man may respond in one dimension to an aspect of the track of the other dimension. The larger remnant term found with more complex control-system dynamics is also compatible with this model. For the strategies which the man uses to compensate for the dynamics play a more dominant part with complex control systems.

(3) *Random variability or "noise."* In some ways the remnant behaves rather like random variability or "noise" injected by the man. Increasing the amplitude of the track, and hence the amplitude of the man's response function, produces an approximately proportional increase in the size of the remnant. The remnant also increases in size as the frequencies in an irregular track are increased. For the man then makes greater efforts to reduce his reaction-time lag by responding to what he predicts the track will do next. The greater the amount of prediction, the greater the errors of prediction. These errors are by definition uncorrelated with the track, and so go to increase the remnant.

Clearly there is something to be said for all three classes of model of the remnant. The remnant may well be as complex as human tracking behavior itself.

The two main weaknesses of the transfer function approach to human tracking are, first, the need to select values of the engineering variables which make the man behave as much as possible like a linear servomechanism. The man is, as it were, put into a strait jacket to minimize behavior which cannot be analyzed by the describing function. The picture which emerges is thus necessarily of an impoverished range of human tracking behavior. However, it is a quantitative picture, in contrast to the more qualitative descriptive approach which has been necessary in this chapter as a result of an absence of sufficient quantitative data in most of the fields referred to. It can be argued with some justification that impoverishment is a small price to pay for quantification. A similar trade-off has been found to be necessary in the field of animal learning (Seward, 1955).

The second main weakness is the size of the remnant term even after this impoverishment. Except for Elkind's (1956) stylus tracking with moderate or low track frequencies, on the average about a quarter of the power in the man's response has to be relegated to the remnant (McRuer & Krendel, 1957, Tables 13 & 14), which by definition contains those aspects of human tracking behavior which do not fit the describing-function model. It may be possible to remedy this weakness in the future by adding further complications to the describing function. Nonlinearity caused by the man's basic intermittency and by his changes in strategy could be dealt with by placing a nonlinear "perfect relay function" in parallel with the main describing function (McRuer & Krendel, 1959). Another possibility is to incorporate hold circuit characteristics in the describing function (Bekey, 1962). Clearly, in the future describing functions are going to be developed and refined until they can fit as much as possible of human tracking behavior.

## The Training Procedure

### PART TRAINING AND TRANSFER

Unlike selective motor learning, in tracking the man may have to learn the nature of any or all of four engineering variables: (1) the track or tracks, and how they are interrelated if at all; (2) the display or displays, and how these are interrelated; (3) the "feel" of the control or controls, and how the directions of control and display movements are related; and (4) the dynamics of the control system or systems. The control of complex vehicles such as spacecraft and high-performance aircraft may necessitate learning in all four areas. In these cases two distinct training problems may arise: It may be possible to train a man on the complete task, but various forms of part training or task simplification may be more efficient; or alternatively the task may be so difficult that the man cannot carry it out at all without some form of pretraining. To date most laboratory experiments have been concerned with the former problem.

For difficult tasks some form of part or progressive-part training is the solution suggested by the vast literature on transfer of training in verbal learning, and by the smaller literature on the learning of motor skills. However, after reviewing this literature Naylor (1962) concluded that with difficult highly organized tasks, where some form of part training is needed, whole training is in fact more efficient. In a difficult but not highly organized compensatory tracking task, part training was found to be the more efficient (Chambers, 1959); the man had to control the movements of a display in three dimensions using three separate rate controls. But when the dynamics of the control system in one dimension of a two-dimensional tracking task interacted with the dynamics in the other dimension, whole training was found to be best (Briggs & Waters, 1958).

Where the component parts of the task interact with each other, task *simplification* has been suggested as an alternative to task "fractionation," as part training has been called. An example of task simplification is reducing the order of a control system from acceleration to rate. A simplification of this kind used early in training on an air interception task which involved a good deal of display learning was found to be as efficient as training on the complex control system representing the dynamics of the aircraft (Briggs, Fitts, & Bahrick, 1958). Simplification is also possible by display aiding or quickening (Briggs, 1961, Exp. 1).

Briggs and Naylor (1962) compared task simplification with part and progressive-part learning on two difficult but not very highly organized compensatory tracking tasks. In both cases the man had to track in three dimensions using a single joystick control. The most complex control system had two or three integrators or exponential time lags in each dimension, while the less complex system had one or two. In neither case did the control dynamics in one dimension interact with those in another. In both cases progressive-part learning was found to be more efficient immediately after transfer than either part learning or task simplification, which involved removing either an integrator or an exponential time lag from each dimension. Whole training was the most efficient method of all, especially with the simpler control system, although the difference from the progressive-part method was unreliable statistically. However, when the man had simply to give orders, instead of carrying out the control task himself, task simplification was found to be more efficient than progressive-part training (Briggs, Naylor, & Fuchs, 1962).

This experiment illustrates the difficulty of attempting to predict even the relative degree of transfer between complex tracking tasks. In any particular case many engineering variables have to be considered, quite apart from procedural variables such as number of training periods and their distribution over time. Many of these variables will almost certainly interact with the variable investigated in the transfer-of-training experiment, and very little is yet known about even the simpler of these interactions. The generalization that more transfer occurs from a difficult version of a task to an easier version than vice versa (Day, 1956) should not be taken to indicate that the training task ought to be more difficult than the criterion task. The degree of transfer may decrease according to the distance of the training task from the criterion task along the dimension under investigation in both the easy and the difficult directions (Muckler, Obermayer, Hanlon, Serio, & Rockway, 1961).

An additional difficulty in comparing part training with task simplification is in optimizing the two training procedures. This involves discovering just when to proceed from one stage of training to the next. It is probably necessary to time the various transitions separately for each person according to his individual progress, so that no time is wasted once a particular aspect of the task has been mastered. Selecting a time which suits the average person may not be precise enough if there are large individual differences. In the case of task simplification, the dynamic trainer may be the answer.

*Dynamic Trainers*

These are teaching machines in which the level of difficulty is adjusted all the time to suit the individual's rate of improvement. Ideally he should always be challenged without being made to feel that the task is too difficult for him. That the optimum is not always easy to hit on in practice is illustrated in an experiment by Briggs (1961, Exp. 1). The aim was to train the man on a second-order or acceleration control system. The control group practiced exclusively with this system, while various experimental groups started with a fully aided or quickened display, and graduated via a rate-aided display to the unaided display. Before changing to a more difficult display either the man's performance had to show that he was ceasing to improve much on his present display, or the man had come to the end of the maximum number of trials permitted on the easier display. Different experimental groups were allocated different criteria of "ceasing to improve much," and a corresponding maximum numbers of trials. Briggs found that there was an optimal training schedule which produced a final level of performance on the unaided display reliably better than an equal time spent tracking exclusively with the unaided display. Training schedules allowing either more or less practice on the easier displays produced a final performance not statistically different from tracking exclusively with the unaided display. The latter schedules can be said to be failures insofar as they represent dynamic trainers. The results thus explain why one dynamic trainer may be said to work well while another is said not to. To produce a successful dynamic trainer it is necessary to hit on just the right relationship between rate of improvement and rate of increase in difficulty.

## DISPLAY AUGMENTATION DURING TRAINING

In scoring time on target, the man may hear his score clicking up every half-second or so on a counter. This *augmented feedback* may have two effects when no summary feedback is given between trials. If the visual display is degraded by visual noise, or as with the Pedestal Sight Manipulation Task the target is fuzzy and the degree of misalignment is consequently not very clear, the man may use the auditory cues to augment the rather poor visual cues, and so do better. In this case, when the augmented feedback is removed after training, performance will return again to its level without augmented feedback (Kinkade, 1963).

The second possible effect of augmented feedback is as a motivator. This occurs when the man already has a display which gives him all the information which he requires in order to perform really well. In this case the augmented feedback produces quicker learning, and a better final level of performance that is maintained when the augmented feedback is subsequently removed (Reynolds & Adams, 1953a). Optimal motivation is obtained by letting the man hear the clicks only when he is performing at or above his average level. He should not be allowed to hear the clicks almost all the time, nor only very occasionally. Since he improves with practice, this means either having a rather strict criterion initially which does not turn out to be too lax after the man has improved, or changing the criterion as the man improves, so that he hears the clicks on the average for the same percentage of each trial. Sudden changes in criterion are detrimental (Kinkade, 1959).

Williams and Briggs (1962) found that the man performed rather better with clicks when he was off target than with clicks when he was on target. Presenting the clicks to the left ear when he was off target to the left, and to the right ear when he was off target to the right, was no improvement over binaural presentation during training, presumably because the man could obtain this directional information anyway from the visual display. On transfer to no augmented feedback, training with the directional monaural feedback gave a reliably worse performance than training with the nondirectional binaural feedback. This was probably because the man had been using the directional monaural information, even though he did not need to so do.

In order to obtain a reliable permanent improvement in performance with augmented feedback over the control condition without, it is necessary to withhold summary feedback at the end of each trial (Briggs, 1961). Summary feedback by itself does in fact produce reliably better performance than auditory augmented feedback by itself (Karlin & Mortimer, 1963).

With a complex control system representing the heading control of an aircraft, augmented feedback only when both the error amplitude and error rate are small is no more effective than no augmented feedback at all. This is the case both for an error amplitude criterion and for an error rate criterion. Augmented feedback only when both the error amplitude and error acceleration are small is also no more effective than no augmented feedback by these two criteria; it is only more effective on an error acceleration criterion (Briggs, 1962c, Exp. 3). Clearly augmented feedback is useful in tracking only when it is made to depend upon the amplitude of the error.

Another form of display augmentation, used by Briggs (1962b) during training on a two-dimensional tracking task representing the dynamics of an aircraft, was to display visual information on error rate and on error acceleration during training, as well as information on error itself. This was found to be no help on finally transferring to displays showing error only. A difficulty with the multiple displays appears to be knowing in what order to look at them.

## DISTRIBUTION OF PRACTICE

In order to produce rapid learning without excessive strain on the man, most experimenters use periods of tracking not exceeding 60 seconds, with rests between of approximately equal duration during which summary feedback can be given. Letting the man track for periods of several minutes on end soon produces a decrement in performance during the trial (Adams, 1952). If the experiment can be spread over days or weeks, there is no need for the man to track for more than 5 or 10 1-minute periods once or twice daily. The work on the distribution of practice has been reviewed by Bilodeau and Bilodeau (1961, pp. 263–268). Insofar as it relates to tracking, it has been carried out almost exclusively on the pursuit rotor using the criterion of time on target, but without summary feedback between trials.

### *Massed vs. Distributed Practice*

Under conditions of massed practice, trials, which may last anything from 10 to 60 seconds, are either continuous or separated by rests of 5 or at most 10 seconds. Under conditions of distributed practice the rests between trials can extend up to 24 hours, but are usually 60 seconds or less. Over the range of 0–30 seconds the shorter the intertrial interval, the lower the standard of performance (Adams, 1954). However, if a series of massed trials is followed by a series of spaced trials, performance soon improves practically to the standard of a group tracking exclusively with spaced trials. Similarly if a series of spaced trials is followed by a series of massed trials, performance soon declines almost to the standard of a group tracking exclusively with massed trials (Reynolds & Adams, 1953b). If a 5-minute rest occurs at the changeover, the performance of the transfer groups soon becomes indistinguishable from the performance of the corresponding no-change groups (Denny, Frisbey, & Weaver, 1955). Thus the effect of massed practice is to depress the standard of performance rather than to depress the rate of learning.

### *Interpolated Rest*

The effect of a rest depends upon whether previously the practice was massed or distributed. After massed practice a rest produces a sudden improvement in performance, the classical reminiscence effect. There is also an initial rapid improvement, the classical warm-up effect. In general the greater the number of massed trials before the rest, and the longer the rest lasts, the greater these effects are (Ammons, 1947). With successive rests the reminiscence effect disappears, but the warm-up effect does not (Adams, 1952). After the warm-up improvement, performance continues to decline as massed practice is continued.

After practice which has been sufficiently distributed, a rest produces a sudden deterioration in performance. This is followed by an initial rapid improvement until performance has almost reached its level before the rest, when there is a more gradual improvement (Adams, 1952). For the rest to have its maximum effect, the man should not be made to watch his partner performing during this time (Adams, 1955).

### *Theoretical Constructs*

Although Ammons (1947) was the first to construct a theory to account for the effects of the distribution of practice, it was Kimble (1949) who imported Hull's (1943) constructs of reactive and conditioned inhibition. Reactive inhibition disappears with rest, just as other drives subside when the aversive stimulus, in this case the continued massed practice, is removed. Some transient intervening variable of this kind is clearly required to describe the mentally fatiguing effect which massed practice has upon performance.

Conditioned inhibition, the habit of resting, results from the reduction of the drive of reactive inhibition by rest. It was introduced to account for the more permanent effects of massed practice which do not disappear after a rest of, for example, 24 hours. However, even these more permanent effects appear to wear off gradually without any specific extinguishing procedure (Adams, 1952). Thus the scientific status of conditioned inhibition in performance on the pursuit rotor is at present somewhat doubtful. The author prefers to think in terms of the challenge effect. Under conditions of massed practice the man soon ceases to perform as well as he can. Even with distributed practice the man eventually settles down to give what he considers to be an adequate performance, which is below what he is capable of. On this view the gradual extinguishing of conditioned inhibition is regarded as the gradual attainment by both massed and distributed groups of a similar standard of performance. It may be due as much to a lowering of the standard of the distributed group as to the raising of the standard of the massed group. If this is so, summary feedback between trials should prevent the development of so-called conditioned inhibition. It might also reduce the amount of decrement attributed to reactive inhibition or mental fatigue.

A difficulty in linking theory to experimental data which is not peculiar either to Hull's (1943) constructs of reactive and conditioned inhibition or to pursuit rotors is the unit of experimental measurement. There is nothing particularly sacred about the size of the target area of the pursuit rotor, yet as the size changes the shape of the learning curve changes nonlinearly (Bahrick, Fitts, & Briggs, 1957). This raises the awkward question: Which particular nonlinear measure of performance is to be equated with the corresponding theoretical construct?

## REFERENCES

Adams, J. A. Warm-up decrement in performance on the pursuit-rotor. *Amer. J. Psychol.*, 1952, **65**, 404–414.

Adams, J. A. Psychomotor performance as a function of intertrial rest interval. *J. exp. Psychol.*, 1954, **48**, 131–133.

Adams, J. A. A source of decrement in psychomotor performance. *J. exp. Psychol.*, 1955, **49**, 390–394.

Ammons, R. B. Acquisition of motor skill. II. Rotary pursuit performance with continuous practice before and after a single rest. *J. exp. Psychol.*, 1947, **37**, 393–411.

Ammons, R. B. Effect of distribution of practice on rotary pursuit hits. *J. exp. Psychol.*, 1951, **41**, 17–22.

Ammons, R.B. Rotary pursuit apparatus. I. Survey of variables. *Psychol. Bull.*, 1955, **52**, 69–76.

Archer, E. J. Effect of distribution of practice on a component skill of rotary pursuit tracking. *J. exp. Psychol.*. 1958. **56**. 427–436

Bahrick, H. P. An analysis of stimulus variables influencing the proprioceptive control of movements. *Psychol. Rev.*, 1957, **64**, 324–328.

Bahrick, H. P., Bennett, W. F., & Fitts, P. M. Accuracy of positioning responses as a function of spring loading in a control. *J. exp. Psychol.*, 1955, **49**, 437–444. (a)

Bahrick, H. P., Fitts, P. M., & Schneider, R. Reproduction of simple movements as a function of factors influencing proprioceptive feedback. *J. exp. Psychol.*, 1955, **49**, 445–454. (b)

Bahrick, H. P., Fitts, P. M., & Briggs, G. E. Learning curves—facts or artifacts? *Psychol. Bull.*, 1957, **54**, 256–268.

Baker, C. H. Factors affecting radar operator efficiency. *J. Inst. Navig.*, 1960, **13**, 148–163.

Battig, W. F., Gregg, L. W., Nagel, E. H., Small, A. M., Jr., & Brogden, W. J. Tracking and frequency of target intermittence. *J. exp. Psychol.*, 1954, **47**, 309–314.

Battig, W. F., Nagel, E. H., & Brogden, W. J. The effects of error-magnification and marker-size on bidimensional compensatory tracking. *Amer. J. Psychol.*, 1955, **68**, 585–594. (a)

Battig, W. F., Voss, J. F., & Brogden, W. J. Effect of frequency of target intermittence upon tracking. *J. exp. Psychol.*, 1955, **49**, 244–248. (b)

Bekey, G. A. The human operator as a sampled-data system. *IRE Trans. on Human Factors in Electronics*, 1962, **3**, 43–51

Bennett, C. A. Sampled-data tracking: Sampling of the operator's output. *J. exp. Psychol.*, 1956, **51**, 429–438.

Bilodeau, E. A. Patterns of internal consistency in multipart skilled performances. *Amer. J. Psychol.*, 1957, **70**, 550–559.

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Birmingham, H. P., & Taylor, F. V. A design philosophy for man-machine control systems. Proc. IRE, 1954, **42**, 1748–1758.

Birmingham, H. P., & Taylor, F. V. A design philosophy for man-machine control systems. In Sinaiko, H. W. (Ed.) *Selected papers on human factors in the design and use of control systems.* New York: Dover Publications, 1961. Pp. 67–87.

Bowen, J. H., & Chernikoff, R. The relationships between magnification and course frequency in compensatory aided tracking. *Nav. Res. Lab. (Washington, D.C.) Rep.* No. NRL-4913, 1957.

Briggs, G. E. On the scheduling of training conditions for the acquisition and transfer of perceptual-motor skills. *NAVTRADEVCEN tech. Rep.*, 1961, No. 836-1.

Briggs, G. E. Pursuit and compensatory modes of information display: A review. *Army Med. Res. Lab. Rep.*, 1962, No. 62-93. (a)

Briggs, G. E. Training for partial-panel control skills. *NAVTRADEVCEN tech. Rep.*, 1962, No. 836-2. (b)

Briggs, G. E. Research on augmented feedback and the acquisition and transfer of skill. *NAVTRADEVCEN tech. Rep.*, 1962, No. 836-3. (c)

Briggs, G. E., & Howell, W. C. On the relative importance of time sharing at central and peripheral levels. *NAVTRADEVCEN tech. Rep.*, 1959, No. 508-2.

Briggs, G. E., & Naylor, J. C. The relative efficiency of several training methods as a function of transfer task complexity. *J. exp. Psychol.*, 1962, **64**, 505–512.

Briggs, G. E., & Waters, L. K. Training and transfer as a function of component interaction. *J. exp. Psychol.*, 1958, **56**, 492–500.

Briggs, G. E., & Wiener, E. L. Fidelity of simulation. 1. Time sharing requirements and control loading as factors in transfer of training. *NAVTRADEVCEN tech. Rep.*, 1959, No. 508-4.

Briggs, G. E., Fitts, P. M., & Bahrick, H. P. Learning and performance in a complex tracking task as a function of visual noise. *J. exp. Psychol.*, 1957, **53**, 379–387. (a)

Briggs, G. E., Fitts, P. M., & Bahrick, H. P. Effects of force and amplitude cues on learning and performance in a complex tracking task. *J. exp. Psychol.*, 1957, **54**, 262–268. (b)

Briggs, G. E., Fitts, P. M., & Bahrick, H. P. Transfer effects from a single to a double integral tracking system. *J. exp. Psychol.*, 1958, **55**, 135–142.

Briggs, G. E., Naylor, J. C., & Fuchs, A. H. Whole versus part training as a function of task dimensions. *NAVTRADEVCEN tech. Rep.*, 1962, No. 950-2.

Brown, I. D., Holmqvist, S. D., & Woodhouse, M. C. A laboratory comparison of tracking with four flight-director displays. *Ergonomics*, 1961, **4**, 229–251.

Calvert, E. S. Manoeuvres to ensure the avoidance of collision. *J. Inst. Navig.*, 1960, **13**, 127–137.

Chambers, R. M. Transfer of training among components of a complex velocity control task. *NADC-MA Lab. Rep.*, 1959, 5920.

Chernikoff, R., & LeMay, M. Effect of various display-control configurations on tracking with identical and different coordinate dynamics. *J. exp. Psychol.*, 1963, **66**, 95–99.

Chernikoff, R., & Taylor, F. V. Reaction time to kinesthetic stimulation resulting from sudden arm displacement. *J. exp. Psychol.*, 1952, **43**, 1–8.

Chernikoff, R., & Taylor, F. V. Effects of course frequency and aided time constant on pursuit and compensatory tracking. *J. exp. Psychol.*, 1957, **53**, 285–292.

Chernikoff, R., Birmingham, H. P., & Taylor, F. V. A comparison of pursuit and compensatory tracking in a simulated aircraft control loop. *J. appl. Psychol.*, 1956, **40**, 47–52.

Conklin, J. E. Effect of control lag on performance in a tracking task. *J. exp. Psychol.*, 1957, **53**, 261–268.

Craig, D. R. Effect of amplitude range on duration of responses to step function displacements. *USAF tech. Rep.*, 1949, No. 5913.

Craik, K. J. W. Theory of the human operator in control systems. 1. The operator as an engineering system. *Brit. J. Psychol.*, 1947, **38**, 56–61.

Craik, K. J. W. Theory of the human operator in control systems. II. Man as an element in a control system. *Brit. J. Psychol.*, 1948, **38**, 142–148.

Davis, R. The limits of the "psychological refractory period." *Quart. J. exp. Psychol.*, 1956, **8**, 24–38.

Day, R. H. Relative task difficulty and transfer of training in skilled performance. *Psychol. Bull.*, 1956, **53**, 160–168.

Denny, M. R., Frisbey, N., & Weaver, J., Jr. Rotary pursuit performance under alternate conditions of distributed and massed practice. *J. exp. Psychol.*, 1955, **49**, 48–54.

Elkind, J. I. Characteristics of simple manual control systems. *Lincoln Lab. Tech. Rep.*, 1956, No. 111. Cambridge, Mass.: M.I.T.

Ellson, D. G., & Gray, F. E. Frequency responses of human operators following a sine wave input. *MCREXD Aero-Med. Lab. Rep.*, 1948, 694-2N.

Ellson, D. G., & Hill, H. The interaction of responses to step function stimuli. I. Opposed steps of constant amplitude. *MCREXD Aero-Med. Lab. Rep.*, 1948, No. 694-2P.

Ellson, D. G., & Wheeler, L. The range effect. *USAF tech. Rep.*, 1949, No. 5813.

Ellson, D. G., Hill, H., & Gray, F. Wavelength and amplitude characteristics of tracking error curves. *TSEAA Aero-Med. Lab. Rep.*, 1947, No. 694-2D.

Fitts, P. M., & Simon, C. W. Some relations between stimulus patterns and performance in a continuous dual-pursuit task. *J. exp. Psychol.*, 1952, **43**, 428–436.

Forbes, T. W. Auditory signals for instrument flying. *J. aeronaut. Sci.*, 1946, **13**, 255–258.

Frost, G. G. A comparison between tracking with "optimum" dynamics and tracking with a simple velocity control. *AMRL-TD Rep.*, 1962, No. 62-150.

Fuchs, A. H. The progression-regression hypothesis in perceptual-motor skill learning. *J. exp. Psychol.*, 1962, **63**, 177–182.

Garvey, W. D. A comparison of the effects of training and secondary tasks on tracking behavior. *J. appl. Psychol.*, 1960, **44**, 370–375.

Garvey, W. D., & Henson, J. B. Interactions between display gain and task-induced stress in manual tracking. *J. appl. Psychol.*, 1959, **43**, 205–208.

Garvey, W. D., & Mitnick, L. L. An analysis of tracking behavior in terms of lead-lag errors. *J. exp. Psychol.*, 1957, **53**, 372–378.

Garvey, W. D., & Taylor, F. V. Interactions among operator variables, system dynamics, and task-induced stress. *J. appl. Psychol.*, 1959, **43**, 79–85.

Garvey, W. D., Sweeny, J. S., & Birmingham, H. P. Differential effects of display lags and control lags on the performance of manual tracking systems. *J. exp. Psychol.*, 1958, **56**, 8–10.

Geldard, F. A. Cutaneous channels of communication. In W. A. Rosenblith (Ed.). *Sensory communication.* New York: Wiley, 1961. Pp. 73–87.

Goldstein, D. A., & Newton, J. M. Transfer of training as a function of task difficulty in a complex control situation. *J. exp. Psychol.*, 1962, **63**, 370–375.

Gottsdanker, R., Broadbent, L., & Van Sant, C. Reaction time to single and to first signals. *J. exp. Psychol.*, 1963, **66**, 163–167.

Gottsdanker, R., Frick, J. W., & Lockard, R. B. Identifying the acceleration of visual targets. *Brit. J. Psychol.*, 1961, **52**, 31–42.

Gould, J. D., & Smith, K. U. Sensory feedback analysis of stereotelevision pursuit tracking. *J. appl. Psychol.*, 1964, **48**, 152–160.

Graham, D., & McRuer, D. *Analysis of nonlinear control systems.* New York: Wiley, 1961.

Hammerton, M. An investigation into the optimal gain of a velocity control system. *Ergonomics*, 1962, **5**, 539–543.

Hammerton, M., & Tickner, A. H. An investigation into the comparative suitability of forearm, hand and thumb controls in acquisition tasks. *Ergonomics*, 1966, **9**, (in press).

Harris, J. D., Pikler, A. G., & Murphy, J. E. Compensatory and pursuit tracking of pitch. *J. acoust. Soc. Amer.*, 1963, **35**, 581–587.

Hartman, B. O., & Fitts, P. M. Relation of stimulus and response amplitude to tracking performance. *J. exp. Psychol.*, 1955, **49**, 82–92.

Heinemann, R. F. D. The relationship between the difficulty level and kind of tracking problem and the type of tracking and type of control. *NAVTRADEVCEN tech. Rep.*, 1961, No. 342-4.

Helson, H. Design of equipment and optimal human operation. *Amer. J. Psychol.*, 1949, **62**, 473–497.

Hick, W. E. Reaction time for the amendment of a response. *Quart. J. exp. Psychol.*, 1949, **1**, 175–179.

Hirsch, J., & Kadushin, I. Experiments in tactile control of flying vehicles. Sixth annual conference on aviation and astronautics, Israel, Feb. 1964.

Hoffeld, D. R., Seidenstein, S., & Brogden, W. J. Proficiency in finger-tracking as a function of number of fingers. *Amer. J. Psychol.*, 1961, **74**, 36–44.

Howell, W. C., & Briggs, G. E. The effects of visual noise and locus of perturbation on tracking performance. *J. exp. Psychol.*, 1959, **58**, 166–173.

Howland, D., & Noble, M. E. The effect of physical constants of a control on tracking performance. *J. exp. Psychol.*, 1953, **46**, 353–360.

Hull, C. L. *Principles of behavior.* New York: Appleton, 1943.

Humphries, M. A Method of estimating "RMS error" in pursuit rotor performance. *Percept. mot. Skills*, 1961, **13**, 211–225.

Hunt, D. P. The effect of the precision of informational feedback on human tracking performance. *Hum. Factors*, 1961, **3**, 77–85.

Jenkins, W. L., & Connor, M. B. Some design factors in making settings on a linear scale. *J. appl. Psychol.*, 1949, **33**, 395–409.

Jenkins, W. L., & Olson, M. W. The use of levers in making settings on a linear scale. *J. appl. Psychol.*, 1952, **36**, 269–271.

Karlin, L., & Mortimer, R. G. Effect of verbal, visual and auditory augmenting cues on learning a complex motor skill. *J. exp. Psychol.*, 1963, **65**, 75–79.

Katz, M. S., & Spragg, S. D. S. Tracking performance as a function of frequency of course illumination. *J. Psychol.*, 1955, **40**, 181–191.

Kelley, C. R. Predictor instruments look into the future. *Control Eng.*, 1962, **9**, 86–90.

Kimble, G. A. An experimental test of a two-factor theory of inhibition. *J. exp. Psychol.*, 1949, **39**, 15–23.

Kinkade, R. G. Augmented feedback and tracking skill. *NAVTRADEVCEN tech. Rep.*, 1959, No. 508-3.

Kinkade, R. G. A differential influence of augmented feedback on learning and on performance. *AMRL-TD Rep.*, 1963, No. 63-12.

Lincoln, R. S. Visual tracking. III. The instrumental dimension of motion in relation to tracking accuracy. *J. appl. Psychol.*, 1953, **37**, 489–493.

McConnell, D., & Shelly, M. W. Tracking performance on a sequence of step functions which approaches a continuous function as a limit. *J. exp. Psychol.*, 1960, **59**, 312–320.

McRuer, D. T., & Krendel, E. S. A dynamic response of human operators. *USAF WADC tech. Rep.*, 1957, No. 56-524.

McRuer, D. T., & Krendel, E. S. The human operator as a servo system element. Parts I and II. *J. Franklin Inst.*, 1959, **267**, 381–404 and 511–536.

Morgan, C. T., Cook, J. S., Chapanis, A., & Lund, M. W. *Human engineering guide to equipment design*. New York: McGraw-Hill, 1963.

Moss, S. M. Tracking with a differential brightness display. I. Acquisition and transfer. *J. appl. Psychol.*, 1964, **48**, 115–122. (a)

Moss, S. M. Tracking with a differential brightness display. II. Peripheral tracking. *J. appl. Psychol.*, 1964, **48**, 249–254. (b)

Mowbray, G. H., & Gebhard, J. W. Man's senses as information channels. In H. W. Sinaiko (Ed.), *Selected papers on human factors in the design and use of control systems*. New York: Dover Publications, 1961. Pp. 115–149.

Mowrer, O. H. Preparatory set (expectancy)—some methods of measurement. *Psychol. Rev. Monogr.*, 1940, **52**, Whole No. 233, 43.

Muckler, F. A., Obermayer, R. W., Hanlon, W. H., Serio, F. R., & Rockway, M. R. Transfer of training with simulated aircraft dynamics. I. Variations in period and damping of the phugoid response. *USAF WADD tech. Rep.*, 1961, No. 60-615 (I). (a)

Muckler, F. A., Obermayer, R. W., Hanlon, W. H., & Serio, F. R. Transfer of training with simulated aircraft dynamics. II. Variations in control gain and phugoid characteristics. *USAF WADD tech. Rep.*, 1961, No. 60-615 (II). (b)

Muckler, F. A., Obermayer, R. W., Hanlon, W. H., & Serio, F. R. Transfer of training with simulated aircraft dynamics. III. Variations in course complexity and amplitude. *USAF WADD tech. Rep.*, 1961, No. 60-615 (III). (c)

Naylor, J. C. Parameters affecting the relative efficiency of part and whole training methods: A review of the literature. *NAVTRADEVCEN tech. Rep.*, 1962, No. 950-1.

Noble, M., Fitts, P. M., & Warren, C. E. The frequency response of skilled subjects in a pursuit tracking task. *J. exp. Psychol.*, 1955, **49**, 249–256.

Obermayer, R. W., Swartz, W. F., & Muckler, F. A. The interaction of information displays with control system dynamics in continuous tracking. *J. appl. Psychol.*, 1961, **45**, 369–375.

Peters, W., & Wenborne, A. A. The time pattern of voluntary movements. *Brit. J. Psychol.*, 1936, **26**, 388–406.

Pikler, A. G., & Harris, J. D. Compensatory and pursuit tracking of loudness. *J. acoust Soc. Amer.*, 1960, **32**, 1129–1133.

Poulton, E. C. Perceptual anticipation and reaction time. *Quart. J. exp. Psychol.*, 1950, **2**, 99–112. (a)

Poulton, E. C. Perceptual anticipating in tracking. *Appl. Psychol. Unit Rep.* No. 118. Cambridge, England, 1950. (b)

Poulton, E. C. Perceptual anticipation in tracking with two-pointer and one-pointer displays. *Brit. J. Psychol.*, 1952, **43**, 222–229. (a)

Poulton, E. C. The basis of perceptual anticipation in tracking. *Brit. J. Psychol.*, 1952, **43**, 295–302. (b)

Poulton, E. C. Eye-hand span in simple serial tasks. *J. exp. Psychol.*, 1954, **47**, 403–410.

Poulton, E. C. The precision of choice reactions. *J. exp. Psychol.*, 1956, **51**, 98–102.

Poulton, E. C. On the stimulus and response in pursuit tracking. *J. exp. Psychol.*, 1957, **53**, 189–194. (a)

Poulton, E. C. Learning the statistical properties of the input in pursuit tracking. *J. exp. Psychol.*, 1957, **54**, 28–32. (b)

Poulton, E. C. On simple methods of scoring tracking error. *Psychol. Bull.*, 1962, **59**, 320–328.

Poulton, E. C. Pursuit tracking with differentiating and integrating control systems. *J. appl. Psychol.*, 1963, **47**, 289–292. (a)

Poulton, E. C. Compensatory tracking with differentiating and integrating control systems. *J. appl. Psychol.*, 1963, **47**, 398–402. (b)

Poulton, E. C. Postview and preview in tracking with complex and simple inputs. *Ergonomics*, 1964, **7**, 257–266.

Poulton, E. C. Tracking a variable rate of movement. *Appl. Psychol. Unit Rep.* Cambridge, England, 1965.

Poulton, E. C., & Freeman, P. R. Unwanted asymmetrical transfer effects with balanced experimental designs. *Appl. Psychol. Unit Rep.* Cambridge, England, 1965.

Poulton, E. C., & Gregory, R. L. Blinking during visual tracking. *Quart. J. exp. Psychol.*, 1952, 4, 57–65.

Regan, J. J. Tracking performance related to display control configurations. *NAVTRADEVCEN tech. Rep.*, 1959, No. 322-1-2.

Reynolds, B., & Adams, J. A. Motor performance as a function of click reinforcement. *J. exp. Psychol.*, 1953, **45**, 315–320. (a)

Reynolds, B., & Adams, J. A. Effect of distribution and shift in distribution of practice within a single training session. *J. exp. Psychol.*, 1953, **46**, 137–145. (b)

Rockway, M. R. The effect of variations in control-display ratio and exponential time delay on tracking performance. *USAF WADC tech. Rep.*, 1954, No. 54-618.

Searle, L. V., & Taylor, F. V. Studies of tracking behavior. I. Rate and time characteristics of simple corrective movements. *J. exp. Psychol.*, 1948, **38**, 615–631.

Seidenstein, S., Chernikoff, R., & Taylor, F. V. The relationship of a retinal-gain index to system performance. *NRL Rep.*, 1960, No. 5548.

Seward, J. P. The constancy of the I-V: A critique of intervening variables. *Psychol. Rev.*, 1955, **62**, 155–168.

Slack, C. W. Learning in simple one-dimensional tracking. *Amer. J. Psychol.*, 1953, **66**, 33–44. (a)

Slack, C. W. Some characteristics of the "range effect." *J. exp. Psychol.*, 1953, **46**, 76–80. (b)

Taylor, F. V., & Birmingham, H. P. Studies of tracking behavior. II. The acceleration pattern of quick manual corrective responses. *J. exp. Psychol.*, 1948, **38**, 783–795.

Taylor, F. V., & Birmingham, H. P. The effect on quick manual corrective responses of interrupting the visual feedback. *Amer. Psychologist*, 1949, 4, 358.

Taylor, F. V., & Birmingham, H. P. Simplifying the pilot's task through display quickening. *J. Aviat. Med.*, 1956, **27**, 27–31.

Thornton, G. B. Comparison of an experimental rolling-ball control and a conventional joystick in speed of tracking on a radar display. *Defense Res. Med. Lab. (Toronto) Rep.* No. 107-1, 1954.

Vince, M. A. The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, 1948, **38**, 149–157. (a)

Vince, M. A. Corrective movements in a pursuit task. *Quart J. exp. Psychol.*, 1948, **1**, 85–103. (b)

Voss, J. F. Effect of target brightness and target speed upon tracking proficiency. *J. exp. Psychol.*, 1955, 49, 237–243.

Warrick, M. J. Effect of transmission-type control lags on tracking accuracy. *USAF tech. Rep.*, 1949, No. 5916.

Welford, A. T. The psychological refractory period and the timing of high-speed performance—a review and a theory. *Brit. J. Psychol.*, 1952, 43, 2–19.

Welford, A. T. Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quart. J. exp. Psychol.*, 1959, 11, 193–210.

Williams, A. C., & Briggs, G. E. On-target versus off-target information in the acquisition of tracking skill. *J. exp. Psychol.*, 1962, 64, 519–525.

# Tracking Behavior
## Comments on Dr. Poulton's Paper

GEORGE E. BRIGGS
*Ohio State University*

Poulton has presented a carefully organized and well-documented dissertation on tracking behavior. His paper starts with input variables, turns next to display variables, then to control device and subsequently machine dynamics variables. It then deals with man as an information processor and concludes with a discussion of training variables. My comments follow this same general sequence of variables, although I do not touch upon all of these categories. Specifically, I found it difficult to add substantially to Poulton's fine review of input variables or the section on control device variables. My remarks, therefore, are directed to the remaining four sections.

## Display Variables

### The Pursuit-Compensatory Dimension

Poulton (1950) provided the first explicit comparison of pursuit and compensatory tracking tasks, and like most of the subsequent research he found superior performance with the pursuit display. There have been several pursuit vs. compensatory comparisons since, and in addition several writers, including Poulton in the preceding paper, have utilized these two display modes as a basis for speaking of two different kinds of tracking tasks. Thus, it is implied that these represent qualitatively different kinds of tasks. Actually, they are more adequately conceived of as the end points on a continuum of "pursuitedness." Senders was the first to demonstrate this (Senders & Cruzen, 1952), but apparently that research, which was published only in an Air Force technical report series, did not receive very wide distribution. It is of some interest, therefore, to review this and some of the subsequent work on the same topic.

Senders pointed out that a 100 percent pursuit display involves zero percent compensatory component and thus, if the target element is displaced, say, four units to the left on the display by a step function input, this is the only movement induced until, of course, the human operator deflects his control device which in turn results in movement of the cursor element. However, suppose we arranged to make the display only 75 percent pursuit; it would then become 25 percent compensatory, and the same step function input would displace the target element only three units to the left but it would displace the cursor simultaneously one unit *to the right*. Thus, system error would still be displayed as a four-unit difference between target and cursor, as in the 100 percent pursuit display, but the input now would affect both rather than only one of the two display elements. As one decreases the percent pursuit component (and thereby increases the percent compensatory component) on a display, the influence of an input decreases with regard to the target element but increases with regard to the cursor until at zero percent pursuit (a 100 percent or pure compensatory) display condition, the input affects only the position of the cursor, the target being fixed in position. There is an interesting illustration of a mixed pursuit-compensatory tracking task from real life: the duck hunter who pursues (actually leads) his target with the gun while compensating for the rocking motion of his flat-bottom boat.

As one might expect, Senders and Cruzen found that tracking accuracy, in terms of time on target, decreased systematically as the percent pursuit component was decreased. Their data did show no significant difference between the 75 and the 100 percent pursuit conditions and, in fact, the latter appeared to be slightly inferior to the 75 percent condition; however, as indicated by Bahrick, Fitts, and Briggs (1957), there are artifacts with dichotomous scores such as time on target and indeed Briggs and Rockway (1966), in a replication of Senders and Cruzen, did find statistically superior performance under a 100 percent pursuit condition to that with a 75 percent pursuit display.

It seems appropriate, therefore, to keep in mind that while we may find it expedient to talk about and build models for two different kinds of tracking tasks, the pursuit and the compensatory, in fact we are dealing with a continuum. And the generality of such theory would be better served if the continuity between pure pursuit and pure compensatory displays was recognized and provided for in the theory.

## PURSUIT VS. COMPENSATORY

In the preceding paper Poulton said, "... it is difficult to understand how a single compensatory display can ever give better results than a single pursuit display." We have two experiments in the literature, both referenced by

Poulton, which obtained superior performance under compensatory tracking conditions: Chernikoff and Taylor (1957) and Obermayer, Swartz, and Muckler (1961). Both found such superiority only with ultra low-frequency inputs and under unburdening conditions.

In terms of "understanding" Chernikoff and Taylor have provided the following, as paraphrased: With ultra low-frequency inputs there are relatively long periods of time when the input is a slow, near ramp function. Thus, the presence of rate or of rate-aided dynamics in the control system permits the operator to match the near-constant rate input with a system output generated by a simple positioning response on the control device. This means that a nearly proportional relationship between system error and the required response exists for relatively long periods of time, and thus the human operator need act primarily as a simple amplifier rather than as a higher-order servo which utilizes error rate and acceleration information in order to generate controlling responses. Birmingham and Taylor (1954) have referred to this as a situation with "stimulus-response integrity": the (compensatory) display contains error amplitude information in a direct form and this is all that is required under rate or rate-aided system dynamics to generate an appropriate output by a simple positioning response.

A pursuit display also provides error information, of course, but it is error information "plus"—plus information from which the operator can estimate the higher derivatives of both the input and output signals. In other words, the pursuit display provides more than that information required to track the ultra low-frequency inputs through rate or rate-aided dynamics, but the compensatory display provides just enough information and it does so in a direct manner, thus its superiority.

Pearl, Simon, and Smith (1955) offered another but more complicated explanation for the superiority of a pure compensatory display mode when the input is in the ultra low-frequency range and rate or rate-aided dynamics are present in the control system. They first defined two basic human output responses in a tracking task: a rapid oscillatory positioning movement and a slower rate-control movement. (These apparently are analogous to the manipulative and travel movements postulated for the discrete cases of motor tasks.) They then pointed out that if most of the tracking error is due to the rapid positioning responses, then rate aiding will improve performance because it will tend to filter out these high-frequency oscillations. They stated that these conditions were found with slow target speeds, thus the beneficial effects of rate aiding. In regard to the superiority of a compensatory display under these same conditions, the authors stated, "The theory . . . also predicts that an aiding device will be of considerable significance in compensatory tracking because . . . (here) . . . rate control motions have been eliminated in large part by zeroing the target" (1955, p. 214).

Thus our "understanding" of the superiority of a compensatory display mode when coupled through the human operator to rate or rate-aided machine dynamics may be summarized as follows:

(1) From Chernikoff and Taylor, one has a higher level of "stimulus-response integrity" under these conditions.

(2) From Pearl, Simon, and Smith, the rate or rate-aided dynamics attenuate the effects of one source of error (the high-frequency positioning movements) while the compensatory display reduces the other source (rate control motions).

The present author trusts that the above discussion will not appear to the reader as a "defense of compensatory displays." They are, after all, superior to pursuit displays only under very limited, but not necessarily unusual circumstances: ultra low-frequency inputs and first- (rate) or higher-order (acceleration, etc.) machine dynamics. And the real importance of such a discussion here is not with regard to the engineering merits of one or the other display mode, but rather to why one is superior to the other and when. In other words, our interest here is with motor skills, and it is most illuminating to note that the reason offered by Chernikoff and Taylor, for example, for compensatory display superiority under the special conditions cited above is that in that mode the information processing requirement placed on the human operator is nearly minimal: he can act as a simple amplifier making only directional and gain adjustments.

## QUICKENING

Poulton's discussion of the quickening principle and of predictor displays is quite brief, and a few additional comments seem warranted. These are, after all, major contributions by our brethren in engineering psychology.

What is the theory behind the quickening principle? The theory is in fact an expression of some motor-skills theory in servo terminology. It looks upon the human operator primarily as an information processor and ignores hypotheses, such as those of Pearl, Simon, and Smith, above, on kinds of motor outputs. Thus it may be incomplete theory, but it is interesting and useful nevertheless.

Birmingham and Taylor (1954) pointed out that if the machine dynamics in a tracking task are second order (acceleration) or higher (jerk, etc.) the human operator is dealing with an inherently unstable system, and he must act as an equalizer in order to bring stability to the system, i.e., to avoid oscillatory outputs of the system which can easily increase in amplitude and thus "go out of control." In theory, then, the human operator must estimate as many derivatives of the displayed system error as there are orders (integral transformations) in the control system, and his output must be based on an appropriately weighted composite of error amplitude, error rate, error acceleration, and as many more derivatives of system error as there are orders in the control system or machine dynamics. For third- and fourth-order machine dynamics this is a prodigious but not an impossible task, and even for a second-order system, such as that encountered in lateral control of an aircraft, this is a task of great difficulty for the novice and of some demand on the skilled pilot. The information processing (derivate estimations, weighting adjustments, and collation of these terms into the appropriate equation or control device movement pattern) required by such tasks is great, but the human operator can perform them successfully; thus, in theory, he is operating as a servo. Some may be offended by this Descartian view of man, but a later section of this paper attempts to show the usefulness of such a view to motor-skills theory.

For the present, then, the Birmingham and Taylor view may be summarized by saying that the relation between displayed error and human output in a second-order system must be

$$\theta_H = K_1 E + K_2 pE + K_3 p^2 E,$$

where $\theta_H$ is the human output, E is system error, $p$ and $p^2$ are the first and second derivatives, respectively, and the *K*'s are weighting constants. Thus, the right-hand portion of the equation represents the information processing required of the human operator to effectively control the system.

Now, with full quickening, the two derivative terms can be deleted from this equation because under quickening the derivatives of system output are sensed, appropriately weighted, and combined with system output itself *prior* to display to the human operator. Thus the derivative information required for stability is already present in the displayed signal and the human need not estimate it. The information processing equation then becomes that of a simple amplifier,

$$\theta_H = K_I E,$$

which says that man merely makes his motor output proportional to the amplitude of displayed error. Of course, he must also transmit directional information to the control device.

Actually, it is possible to even further reduce the information processing demands on the human tracker. If one employs quickening plus display aiding,[1] then the tracker need transmit only directional information to the control device, i.e., he need act only as a simple relay!

---

[1] Display aiding, in this case, requires taking the first derivative of what would be displayed error in full quickening and adding this to such an error signal prior to display to the human operator. This has been identified as superquickening by Rund, Birmingham, Tipton, and Garvey (1957).

Poulton is rather impatient with the quickening principle, stating, "If the display has been fully quickened, the task does not in fact require a man at all; in flying an aeroplane the autopilot can be used." One could add now: How ludicrous it would be to require a human operator in a superquickened system, unless, of course, one wanted to provide gainful employment for the severely brain damaged! Actually, this is not the point. The quickening principle is of use to us here not as an engineering device but as a concept which helps develop theory and thus obtain a level of understanding of motor-skill behavior: namely, the view of man in a tracking task as an information processor with servolike functions of derivative estimation, weighting, and collation of signals.

I cannot resist defending quickening and especially the predictor display as devices at this point, despite the fact that we are interested here in skill, not engineering. Both display techniques do permit the human operator to time-share effectively among several concurrent activities. And in the case of the predictor display, the operator is freed to a large degree from the demanding servo role and permitted to work a greater portion of his time on the more lofty plane of the strategy maker: Just how will I level the submarine out at 250 feet? In an underdamped, overdamped, or critically damped manner? If underdamped, how much overshoot will I permit and therefore how long may it take to attenuate the oscillations around my desired depth, etc.?

## The Control System

### Lags

At this point I should like to skip over the section of Dr. Poulton's paper on The Control and provide some remarks to the section titled The Control System. Specifically, I should like to provide data from and comment on some research with transmission, exponential, and sigmoid lags.

In the first of a series of studies, Warrick (1949) required his subjects to track a complex sinusoidal input through a positional control system but with transmission lags of 0, 40, 80, 160, and 320 milliseconds. He found a linear deterioration in time on target as lag was increased, and even performance at 40 milliseconds of lag was statistically inferior to that at 0 millisecond, the lag-free condition. In a second study Warrick (1955) studied exponential lags between 0 and 1000 milliseconds. He used a step function input and obtained results similar to those cited above except that an exponential rather than a linear deterioration in performance was noted.

Levine (1953) used both longer (2.7 seconds) and shorter (15 milliseconds) exponential lags and the four subjects tracked an input composed of 3, 5, and

11 cycles per second sinusoids. The results were similar to Warrick's except that for delay times in excess of 150 milliseconds performance deterioration was linear rather than log-linear. This consistency across three experiments would suggest the conclusion that any lag, even one so short as 15 milliseconds, is detrimental to tracking performance. As is usually the case for behavioral scientists, such a simple conclusion is erroneous and a definitive experiment by Rockway (1954) provides us with a more important conclusion.

Rockway varied both exponential lag time constants and system gain. The control-display (C/D) ratios from least to most sensitive system conditions were 1/3, 1/6, 1/15, and 1/30; and the four lag times were 0.3, 0.6, 1.5, and 3.0 seconds. The subjects tracked the same input as in the Levine study in a two- rather than a single-dimensional tracking task. Only with the least sensitive system did performance deteriorate monotonically, as in the other studies reported above, with a C/D ratio of 1/6, optimum performance occurred at a 0.6-second lag time; at the C/D ratio 1/15 superior performance occurred at the 1.5-second lag; and at the 1/30 C/D ratio performance was best at a 3.0-second lag.

Therefore it would be erroneous to conclude that any system lag is detrimental to tracking accuracy. Instead, the more appropriate conclusion is that whereas in low gain systems an exponential lag may be very detrimental, performance with that same amount of lag may well be excellent, and even better than no lag at all, in systems of higher sensitivity.

There appear to be two reasons for the superiority of tracking through an exponential lag in high gain systems. For one, given a particular lag and a low gain system, there is a delay in information feedback on the control responses. However, as system gain is increased, this increases the slope of the initial, linear portion of the output of an exponential lag filter. Therefore, while a high gain system does not alter the time required for the filter output to asymptote (compared with that in a low gain system), the higher gain does produce a given amplitude of the initial output at an earlier time than does a lower gain, and thus a decrease in the delay of feedback.

A second, possible reason for the superiority of tracking through lags with high gain systems is based on the nature of the shape of the output of an exponential filter to a step function input: it is linear just like the output of a rate system. Thus, since Rockway used an input in the ultra low-frequency range, there were times when that input exhibited constant rate or ramp characteristics, and like the Chernikoff and Taylor (1957) study cited earlier, we have here a tracking task with fairly high "stimulus-response integrity," i.e., by using step-function control movements, or even dither, the operator could treat the exponential lag as if it were an integral lag (rate dynamics) and thereby take advantage of this "unburdening" and the S-R integrity present in tracking inputs of near-constant velocity.

Rockway provided graphic recordings as examples of tracking responses under several of the experimental conditions. It does appear that his subjects employed step function control movements quite extensively, especially with the longer lags and the more highly sensitive systems. Thus, the subjects apparently behaved as required for validity of the "unburdening" explanation, and it is reasonable that both explanations are valid for these data.

### EXPONENTIAL VS. SIGMOID LAGS

Conklin's dissertation (1957) has been summarized by Poulton, who pointed out that performance with sigmoid lags was uniformly inferior to that with exponential lags of comparable time constants, except for random input signals where performance was equally poor for both types of lag.

It is interesting to note that in most of Conklin's data, performance with a 0.25-second sigmoid lag was comparable to that with a 1.0-second exponential lag, and one may be curious as to why this comparability of performance for such different lag time constants. After all, a sigmoid lag with a 0.25-second time constant approaches output asymptote four times more quickly that does an exponential lag with a 1.0-second time constant, and thus the "full" revelation of the effects of a control movement in the latter case is considerably delayed over that in the former. One can only conclude that it is not the time required to obtain *complete* knowledge of results that is important. What is important here, then?

I speculate that it is the very initial level of output of the two types of lag that determines performance most directly; specifically, the first *third* of the output of the lag filter appears to define the critical period. This supposition is arrived at by noting that an output plot of an exponential lag of 1.0 second and of a sigmoid lag of 0.25 second shows that the two functions are fairly comparable up to one third of the total output pattern, but after that, the sigmoid lag output rapidly becomes greater than that of the exponential lag. Since Conklin found comparable performance for these two lag conditions, the most simple inference is that this was determined by the comparability of the very initial output characteristics of the two filters.

If this speculation is valid, it points out how critical the first few fractions of a second are to performance through machine lags. What makes this even more remarkable is that Warrick (1949) stated that with a 40-millisecond lag the subject's performance was impaired even though no *perceptible* lag was present!

### DITHER

Before turning to the final set of comments on Poulton's fine paper, I would like to amplify briefly on the dither mode of responding. This is a fascinating kind of response but has never been studied systematically, to my knowledge. Dither is defined as a high-frequency, low-amplitude oscillation of the control device. It is found when the machine dynamics involve lags and it is most apt to occur when the subject is highly motivated to keep error to a minimum. A common example is found in driving over icy roads: the driver tends to rotate the steering wheel rapidly back and forth in small amplitude oscillations.

Dither produces two desirable effects: it increases the amount of information feedback (do I still have control?) and it tends to reduce complex machine dynamics (second and third order or higher) to a ratelike or first order of control. The latter is based on the same reasoning as that used in the comment above on Rockway's data: by using small amplitude and alternating step function control movements, the operator limits the system output to the initial, nearly linear portion of the total output pattern. Thus, by using dither, the operator can make a higher-order system behave like a simple first-order system.

A rather interesting example of the effectiveness of the dither mode appeared in the study by Levine (1953), cited above. One of the four operators employed in that study actually exceeded his proficiency under the lag-free condition on lags of 60 and 90 milliseconds; the other three "normal" operators did more poorly on these lags, of course. As we might now expect, Levine noted that the unusual operator, Subject D, ". . . kept the knob in continual movement by turning (it) back and forth through approximately 20° at a very high frequency" (1953, p. 13). Levine excluded Subject D from his analyses of the data, but although he may have felt some chagrin at being labeled "anomalous," Subject D should be heartened by the realization that he did perform at a more sophisticated level than did his fellow operators, Subjects A, B, and C!

## The Human Operator

Turning now to Poulton's section, The Man, I should like to relate three experiments mentioned in that section by reference to some theory first proposed by Fitts, Bahrick, Noble, and Briggs (1959, pp. 12.47–12.50). Specifically, the research by Garvey and Mitnick (1957), by Garvey (1960), and by Fuchs (1962) each dealt with changes in the human transfer function "constants" as the operators practiced and/or were subjected to task-induced stress, and these three studies support, I believe, the progression-regression hypotheses of skilled behavior in continuous (tracking) tasks.

The progression hypothesis was developed during preparation of the Fitts *et al.* (1959) report on the basis of Garvey and Mitnick's (1957) data and on the basis of some earlier results from the Goodyear Laboratories (Goodyear Aircraft Corp., 1955). In both cases the order of the transfer function was higher for the well-practiced operator than for either the same or another operator with less training. By "order" of the transfer function is meant the number of derivative terms needed in the differential equation used to define the basic transfer function equation, and by "needed" I mean the number of derivative terms for which a transfer function analysis generated significant constants.

The hypothesis states that a skill novice will attempt to generate a control response, $\theta_H$, by responding directly and proportionately to the amplitude of system error:

$$\theta_H = K_1 E,$$

where the terms are as defined earlier, i.e., the novice will try to respond as a simple amplifier and merely transmit the displayed amplitude information to the control device. This may be defined as the zero order of control, and unless the input is only a step function, the operator will observe considerable system error. The hypothesis then states that the operator will, with this experience, progress to the first order of control and generate responses based on not only the amplitude but also the velocity characteristics of the error signal,

$$\theta_H = K_1 E + K_2 pE,$$

and further, that with more practice he will progress to a second order of control wherein amplitude, rate, and acceleration of system error will enter into a determination of his controlling responses:

$$\theta_H = K_1 E + K_2 pE + K_3 p^2 E.$$

Theoretically there is no upper limit to the order of control which the operator may acquire, although in real life, few system dynamics and/or inputs require anything higher than the second or acceleration order of control.

The Garvey and Mitnick (1957) data, a selection of pilots from the Goodyear research (see Fuchs, 1962, p. 181), and the initial acquisition data of Fuchs all support the progression hypothesis; however, neither the Fuchs nor the Garvey and Mitnick subjects seem to have started at the zero or amplitude level of control; rather, both sets of subjects appear to have started at the velocity level of control. In this regard, it is not unreasonable to assume that after 18 or more years of experience with a variety of vehicles, these subjects could progress so rapidly from a zero to a first order of control that the summarized results would not reveal the earlier, more primitive level. Perhaps the best place to observe such a transition would be with very young children in tricycle control tasks.

The regression hypothesis states simply that having attained some level of control, skill if you will, the operator will regress to a more primitive level under stress, but that he can re-progress under continuing exposure to stress back toward a more sophisticated level of control. Both the Garvey (1960) and the Fuchs (1962) data support this hypothesis. In particular, Fuchs' results show an initial regression to a strong emphasis on amplitude control with subsequent recovery to a velocity order of control; the Garvey experiment was designed to show only regression, not the potential recovery phase.

It appears, therefore, that the progression-regression hypotheses are adequately supported by the literature, although recovery from stress effects could be more completely explored. While this may not represent a very heady level of behavioral theory, it does represent a more direct route to theory in motor skills than did early attempts to translate S-R theory from the discrete case, where it was developed and was useful, to the continuous case where semi-continuous behavior encourages concepts and analytic methods of measurement which treat the continuities not the discontinuities in responding.

## Training Procedures

To conclude the comments of Poulton's paper, I would like to take exception to his interpretation of a series of studies on the question of part vs. whole training which was conducted at Ohio State several years ago. Specifically, Poulton reviews the first two of a three-experiment series by Briggs, Naylor, and Fuchs (1962) and then states with regard to the second that, "This experiment illustrates the difficulty of attempting to predict even the relative degree of transfer between complex tracking tasks." Actually, the most important study was the third in that series, and while it was not a tracking study, there is no apparent reason that the concepts and results do not apply to tracking behavior. That study indicates that while one might be hard pressed to predict a percent transfer figure for a particular training procedure, it is possible to predict the presence *and* the pattern of a second-order interaction with training method (part vs. whole) as one of the variables and transfer task complexity and organization as the other two.

Specifically, the design of the third study, which may be read in a more available publication by Naylor and Briggs (1963), was based on a review of the literature by Naylor (1962) and on such research as that of Briggs and Waters (1958) and Seymour (1956). The experimental prediction was that, ". . . there should be an interaction between task complexity, task organization, and training methods such that (a) for tasks of relatively high organization as task complexity is increased, whole-task training should become relatively more efficient than part-task methods; however, (b) for tasks of relatively low organization (all task dimensions being independent) an increase in task complexity should result in a part-task method becoming superior to the whole-task method" (Naylor & Briggs, 1963, p. 218). The results supported the predictions, and the latter truly were made *a priori*, as shown by Table I.

TABLE I
TRANSFER INDICES[a]

| Task organization | Task complexity | |
|---|---|---|
| | Low | High |
| Low | 27.2% | 233.3% |
| High | 67.7% | −20.0% |

[a] From Naylor and Briggs (1963).

The index used was $(C_i - E_t)/(C_i - C_t) \times 100$, where $C_i$ and $C_t$ were the initial and transfer trial performance levels of the control group (whole training), respectively, while $E_t$ was the performance of an experimental group (progressive-part training) upon transfer to the whole task. Thus, a transfer index less than 100% indicates a superiority for whole training while one greater than 100% indicates superiority for the progressive-part training method.

Obviously, the predicted interaction among training method, task organization, and task complexity was obtained, and more importantly, the pattern of the interaction predicted *a priori* was obtained also. So I cannot agree with Poulton's conclusion that one necessarily has difficulty predicting relative transfer. We were either lucky or we read the previous research properly and made some appropriate extrapolations; I prefer the latter conclusion, for obvious reasons.

As indicated earlier, these results were not generated in a tracking task, but again, there is no apparent reason that they would not obtain in that kind of behavioral task. What is needed here is an adequate and general means to quantify task organization: the degree of cross coupling or interrelationship between the several dimensions of a task. The matter of defining task complexity in terms of the number and kind of information processing steps required of man appears adequate to specify this variable but task organization is more illusive.

## Concluding Remarks

I conclude this commentary with a few remarks on the matter of the continuous vs. the discrete in motor-skill behavior. It is apparent that man's behavior in a variety of tasks, such as those found in studies of verbal learning, concept formation, retention of verbal habits, conditioning and extinction, motivation, etc., can be adequately measured by discrete scoring procedures and it is equally obvious that theory for these discrete cases can be interesting and useful. I question, however, that we do justice to motor-skill theory by emphasizing the discrete rather than the continuous character of the behavior. Of course, one can demonstrate discontinuity in motor outputs: Vince (1948) did so to define the psychological refractory period, Woodworth did in his doctoral dissertation (1899) to explain simple positioning responses, and so on, but while these are useful they should not overshadow the point that in tracking continuous inputs through machine dynamics, man exerts a *pattern of force* on the control device, and it is this continuous output that should be the main focus of our attention. Since man can and does produce a time-varying, graded response in tracking, it is important that theory of tracking performance utilize concepts appropriate to the continuous case.

This is not to deny that man in tracking even continuous inputs does not exhibit discontinuities in his output; indeed, the earlier discussions of the efficacy of step function control movements and dither illustrate the presence and importance of such discrete responses. Further, most writers have agreed that the human operator is intermittent as a processor of information even in continuous tracking tasks, and it has been suggested by Fitts *et al.* (1959) that the sampling servo is a more appropriate model for the human being than is the continuous servo model.

But again, intermittent or not, in the tracking task one is dealing with force patterns over time, and these continuities deserve our first consideration.

## REFERENCES

Bahrick, H. P., Fitts, P. M., & Briggs, G. E. Learning curves – facts or artifacts? *Psychol. Bull.*, 1957, 54, 256–268.

Birmingham, H. P., & Taylor, F. V. A human engineering approach to the design of man-operated continuous control systems. *USN Res. Lab. Rep.*, 1954, No. 4333.

Briggs, G. E., & Rockway, M. R. Learning and performance as a function of the percent pursuit component in a tracking display. *J. exp. Psychol.*, 1966 (in press).

Briggs, G. E., & Waters, L. K. Training and transfer as a function of component interaction. *J. exp. Psychol.*, 1958, **56**, 492–500.

Briggs, G. E., Naylor, J. C., & Fuchs, A. H. Whole versus part training as a function of task dimensions. *NAVTRADEVCEN Rep.*, 1962, No. 950-2.

Chernikoff, R., & Taylor, F. V. Effects of course frequency and aided time constant on pursuit and compensatory tracking. *J. exp. Psychol.*, 1957, **53**, 285–292.

Conklin, J. E. Effect of control lag on performance in a tracking task. *J. exp. Psychol.*, 1957, **53**, 261–268.

Fitts, P. M., Bahrick, H. P., Noble, M. E., & Briggs, G. E. Skilled performance. *USAF WADC final Rep.*, 1959.

Fuchs, A. H. The progression-regression hypothesis in perceptual-motor skill learning. *J. exp. Psychol.*, 1962, **63**, 177–182.

Garvey, W. D. A comparison of effects of training and secondary tasks on tracking behavior. *J. appl. Psychol.*, 1960, **44**, 370–375.

Garvey, W. D., & Mitnick, L. L. An analysis of tracking behavior in terms of lead-lag errors. *J. exp. Psychol.*, 1957, **53**, 372–378.

Goodyear Aircraft Corp. Investigation of control "feel" effects on the dynamics of a piloted aircraft system. *Goodyear Aircraft Corp. Rep.* No. GER-6726 (1955).

Levine, M. Tracking performance as a function of exponential delay between control and display. *Wright Air Develpm. Cent. Tech. Rep.* No. WADC-53-236 (1953).

Naylor, J. C. Parameters affecting the relative efficiency of part and whole practice methods: A review of the literature. *NAVTRADEVCEN Rep.*, 1962, No. 950-1.

Naylor, J. C., & Briggs, G. E. Effects of task complexity and task organization on the relative efficiency of part and whole training methods. *J. exp. Psychol.*, 1963, **65**, 217–224.

Obermayer, R. W., Swartz, W. F., & Muckler, F. A. The interaction of information displays with control system dynamics in continuous tracking. *J. appl. Psychol.*, 1961, **45**, 369–375.

Pearl, B. E., Simon, J. R., & Smith, K. U. Visual tracking, IV: interrelations of target speed and aided-tracking ratio in defining tracking accuracy. *J. appl. Psychol.*, 1955, **39**, 209–214.

Poulton, E. C. Perceptual anticipation in tracking. *Appl. Psychol. Unit (Cambridge) Rep.* No. 118 (1950).

Rockway, M. R. The effect of variations in control-display ratio and exponential time delay on tracking performance. *USAF WADC tech Rep.*, 1954, No. 54-618.

Rund, P. A., Birmingham, H. P., Tipton, C. L., & Garvey, W. D. The utility of quickening techniques in improving tracking performance with a binary display. *USN Res. Lab. Rep.* No. 5013 (1957).

Senders, J. W., & Cruzen, M. Tracking performance on combined compensatory and pursuit tasks. *USAF WADC tech. Rep.*, 1952, No. 52-39.

Seymour, W. D. Experiments on the acquisition ot industrial skills. Part 3. *Occup. Psychol.*, 1956, **30**, 94–104.

Vince, M. A. The intermittency of control movements and the psychological refractory period. *Brit. J. Psychol.*, 1948, **38**, 149–157.

Warrick, M. J. Effect of transmission-type control lags on tracking accuracy. *USAF Tech. Rep.* No. 5916 (1949).

Warrick, M. J. Effect of exponential-type control lags on speed and accuracy of positioning of a visual indicator. *USAF WADC tech. Note*, 1955, No. 55-348.

Woodworth, R. S. The accuracy of voluntary movement. *Psychol. Monogr.*, 1899, **3**, No. 2 (Whole No. 13).

## Chapter 9

# Cybernetic Theory and Analysis of Learning[1]

KARL U. SMITH[2]
*Behavioral Cybernetics Laboratory, University of Wisconsin*

The main objective of this chapter is to define the role of feedback-controlled motor activity in determining the course of learning. This constitutes a cybernetic approach to the understanding of performance during learning—an approach which recognizes as determinants of learning the feedback factors which govern various aspects of response and of receptor and neural input functions. The cybernetic approach to performance and learning is supported by new laboratory techniques, to which we have given the name *experimental behavioral cybernetics*, and is generating a new technical language as well as theoretical ideas which challenge many of the formulations of current learning theory.

The basic assumption of behavioral cybernetics is that the organization of motor performance and the course of learning are related to the properties of closed-loop feedback control which an animal or individual maintains over its own behavior. The methods of feedback analysis which we have developed utilized electronic, optical, magnetic recording, closed-circuit television, and closed-loop computer systems and have been applied to many different problems of response regulation and organization. The guiding principle in such cybernetic research is to measure the effects of varying

---

[1] The experiments described in this chapter were carried out in a project on Sensory Feedback Analysis of Behavior, National Science Foundation, and in a project on Geometric Organization of Human Motion, Project M-4469, National Institutes of Mental Health.

---

[2] The experimental cybernetic computer research described in this chapter was conducted with the collaboration of Sherman Ansell and Terrence Schuh, who aided in developing the closed-loop computer system. Electrical engineering assistance was obtained from Gerald Servos and Harold Washington, Elmhurst, Illinois.

feedback properties on the efficiency of performance and the characteristics of learning.

## Cybernetic Approach to Learning

The experimental foundation of the cybernetic interpretation of behavior can be traced to the instrumental research of psychologists on radar and gun tracking systems during World War II. Subsequently, Wiener (1948) coined the word "cybernetics" to refer to the properties of directional guidance of man-governed systems and described certain mathematical characteristics of such systems. However, Wiener's ideas deflected the newly named discipline from its experimental base in wartime psychological research and oriented it toward the areas of computer simulation, information theory, and mathematical modeling of psychological and physiological systems (Chase, 1963; Fairbanks, 1954; Fitts, 1951).

In the cybernetic research to be described here, the emphasis is on experimental analysis of feedback interactions in closed-loop self-governing systems (K. U. Smith, 1961a,b, 1962, 1963a,b, 1964; K. U. Smith & W. M. Smith, 1962; K. U. Smith & M. F. Smith, 1966). A fundamental distinction is made between closed-loop and open-loop operating systems. Closed-loop systems are self-regulating and their characteristics of performance and learning reflect the properties of closed-system feedback control. Open-loop systems or operations are limited to those in which reactions are direct functions of external or extrinsic stimulating or information conditions uninflenced by internal or intrinsic feedback regulation. Closed-loop operations are analyzed in terms of direct functional relationships between receptor, neural, and motor action or between stimulus and response, whereas open-loop operations are studied on a statistical or probabilistic basis, utilizing generic specifications of behavior. Closed-loop systems function in terms of response-generated sensory signals and their organization is based primarily on built-in properties of the sensorimotor process, whereas traditional interpretations of behavior assume one-way linear processes relating external events with behavioral effects. Inasmuch as conventional reinforcement and conditioning theories describe learning as being determined by extrinsic open-loop contingencies of reward and punishment or temporal association, they are incompatible with the assumptions guiding experimental cybernetic research.

We specify the following general properties of cybernetic systems: (1) they are self-governed; (2) they operate in terms of reciprocal closed-loop circuits of activity composed of dynamic response and sensing processes; (3) they involve continuous control of such closed-loop circuits of activity; (4)

they are directly guided in both sensing and response in terms of feedback control; (5) different component activities of the closed-loop system are integrated through feedback processes; (6) the feedback control of such response systems can be transformed by instruments and communication (symbolic) systems of specialized design.

Cybernetic theory assumes that the efficiency of performance and the course of learning depend on properties of the feedback-control process, which can vary in a number of ways according to the design of the cybernetic system in relation to the design of its physical environment. These variations include (1) introduction of delay or time lag in the feedback loop; (2) geometric displacement of sensory feedback with respect to the pattern of movement; (3) time sampling or intermittency of feedback; (4) temporal inversions of time samples; (5) kinetic modulations of feedback; (6) complexity of the feedback pattern as defined by the number of component movement systems involved in a response; (7) precision of feedback control; and (8) form and complexity of the instrumental and symbolic transformations involved in a closed-loop system. The assumption is that these intrinsic characteristics of the feedback loop, rather than open-loop reinforcement and associative contingencies, determine the fundamental properties of learning functions. It also is assumed that special properties of development, memory, generalization, transfer, and inhibition are related to variations in the pattern of feedback control.

## Mechanisms of Feedback Control

In applying these cybernetic concepts to an analysis of behavior, we make several assumptions about the nature of the response mechanisms of behaving systems, namely: (1) that the feedback process depends primarily on neural detection of spatial differences in stimulation, i.e., detection of angular displacements between movement and its feedback patterns; (2) that response (and thus its feedback-control pattern) is always multidimensional and involves the integration of tremor reactions and postural, transport, and manipulative movements; (3) that pattern perception is regulated by means of dynamic control of receptors by the movement systems; (4) that somatic movement is controlled spatially in two coordinate dimensions corresponding to the vertical and bilateral axes of the body by means of reference feedback mechanisms of postural and transport movements; (5) that the effectiveness of tools and symbol systems in human control patterns depends on how and to what extent they transform basic feedback properties; and (6) that the nervous system functions in cybernetic control by integrating feedback signals generated by both internal and external behavior.

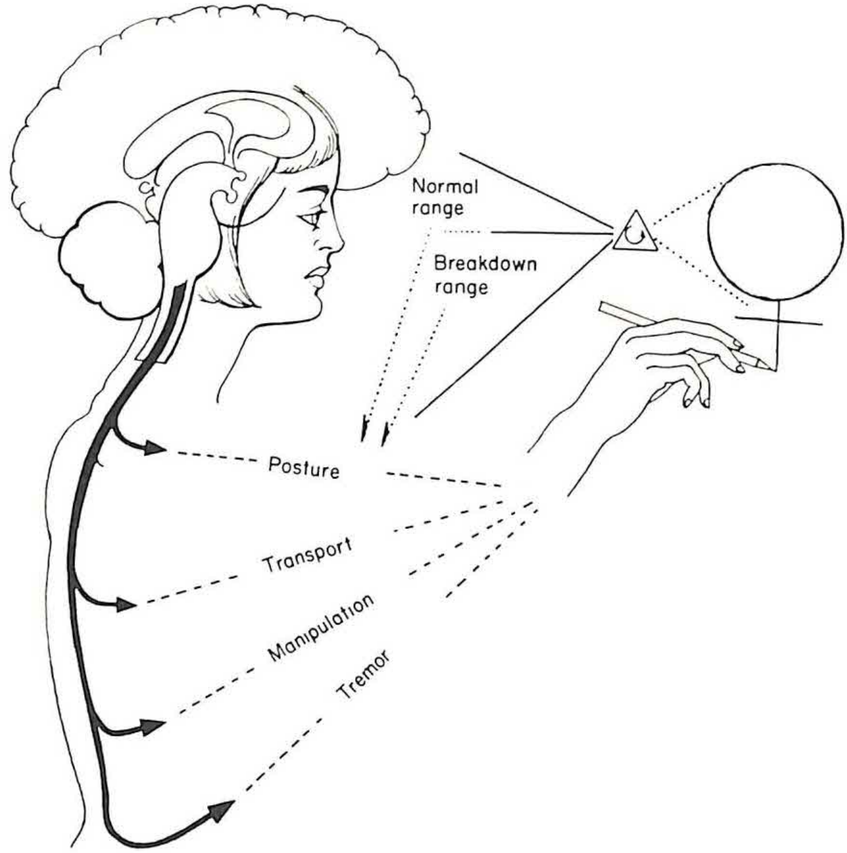


FIG. 1. The cybernetic concept of neurobehavioral regulation of response. Normal and breakdown ranges of angular displacement of feedback involved in regulating direction of response. Differentiation of the postural, transport, and manipulative subsystems of response. The postural and transport systems serve as generalized inertial directional guidance systems for maintaining feedback orientation of all movements with respect to the body axes.

Figure 1 illustrates the first of these points, that the detection of angular displacement of movement-generated feedback is basic to all levels of cybernetic regulation and learning. This figure introduces the concepts of "normal" and "breakdown" ranges of angular displacement of feedback. For any given movement, the feedback pattern can be displaced spatially through a normal range without affecting the efficiency of its performance or learning. Displacement beyond this normal range, however, causes disturbances or breakdowns in response control and learning. Within the breakdown range, the degree of disorganization or learning disturbance varies as a function of the angular displacement.

As indicated in Fig. 1, we assume that response regulation of feedback is multidimensional and involves separate but interrelated tremor, postural, transport, and manipulative mechanisms. The gravity-controlled postural mechanism is the most generalized of these four systems. The bilateral transport motion system is differentiated from the postural mechanisms and is regulated primarily by neuromotor mechanisms that detect bilateral differences in movement-produced feedback stimuli. The third motion subsystem includes the refined articulated activities of the face, hands, feet, and mobile receptor systems such as the eye. Tremor movements generated by neural and ballistic effects of cardiac action along with other movements serve to keep the receptors of the body sensitive to stimulation by preventing adaptation. We believe that these different behavior subsystems are controlled differentially and are integrated into unified response patterns by means of specialized neural feedback mechanisms. We also believe that there are distinct feedback mechanisms which maintain the sensitivity of exteroceptive receptors, determine their orientation in space in relation to body movements, and process or control the stimulus patterns imposed on the receptor surfaces.

## Cybernetic Analysis of Stimulus-Response Interaction in Learning

The feedback concept suggests a new interpretation of the stimulus-response (S-R) interaction in learning. Following Thorndike's (1913) and Watson's (1924) early writings, learning psychologists generally have used the generic concepts of stimulus and response (Skinner, 1935), but have devised no methods or experiments to test their assumptions about the characteristics of S-R relationships. In contrast, experimental cybernetics analyzes the intrinsic S-R interactions of the primary responses or reflex patterns which are modified by learning. The response in a learning situation is considered not as a relatively invariant unit of behavior, but as a variable control pattern which can be analyzed along many parameters. Thus, the reflex is conceptualized not as an S-R unit but as a pattern of feedback-mediated interactions between the motor and sensory systems which provide a measure of behavioral control.

The cybernetic theory proposed here suggests that the brain be conceptualized not as a synaptic system but as a neuronic system, in which internuncial cells function as direction-specific and stimulus-specific units in detecting differences in movement-produced stimulation. Learning would involve imparting directional stimulus specificity to neurons whose specificity is not predetermined by intrinsic factors of development. Presumably, the specificity of the internuncial nerve cells is altered by the effects of their own internal closed-loop processes in response to presynaptic stimulation. Neurons thus would become responsive only to certain loci or directions of presynaptic stimulation, or have a lowered threshold of reaction to such directions of stimulation. In assuming that the nerve cells themselves rather than the synapses become specialized in development and learning, this view aligns itself with the classical doctrine of specific nerve energies and with the cellular theory of biological organization, development, and evolution. It is our belief that the doctrine of synaptic integration and the related hypotheses of temporal association violate both the principle of specific nerve energies and the central tenets of biological cellular theory by assuming that response is determined by the electrical or chemical environment external to the nerve cells. In contrast, the neuronic interpretation of brain function integrates the concepts of both learning and reflex action with the central theoretical constructs of biology, and makes possible a unified theory of evolution and development.

## EXPERIMENTAL CYBERNETIC TECHNIQUES

We distinguish among several areas of feedback research which utilize distinctive experimental methods, namely: (1) receptor cybernetics or the study of feedback mechanisms governing characteristics of the input afferent process; (2) studies of sensorimotor feedback control; (3) studies of psychophysiological control or vital interactions among the biochemical, physiological, and behavioral levels of the response mechanism; (4) neurocybernetics or the study of response regulation of brain mechanisms; and (5) analyses of transformations of feedback control in instrumental and symbolic behavior. The study of cybernetic transformations is relevant on the one hand to basic human engineering research on the design of tools and machines, and on the other to the analysis of symbolic design factors in language and communication.

The primary methodological requirement for cybernetic research is to devise techniques which close the sensorimotor circuit of a particular response system so as to introduce experimental variations of the controlling feedback process and to test their effects on performance and learning. The historical background of this type of research encompasses human engineering research on tracking in World War II (Gebhard, 1948; Phillips, 1943; Warrick, 1949), the classical studies of experimental displacement of vision by means of lenses, prisms, and mirrors (Brown, 1928; Ewert, 1930; Siipola, 1935; Stratton, 1896, 1897, 1899; Wooster, 1923), and the more recent observations on the effects of delayed hearing on sound-producing activities (Black, 1951, 1955; Fairbanks, 1955; Fairbanks & Guttman, 1958; Lee, 1950a, b, 1951). That these different kinds of sensorimotor study constitute a distinct area of feedback research was indicated originally in reports of research on televised displacement of the visual image of motion (K. U. Smith, 1961a, b; W. M. Smith, Smith, Stanley, & Harley, 1956).

Figure 2 illustrates four laboratory techniques for cybernetic research. These include utilizing an optical system to invert the visual feedback of motion, using an audio-tape system to delay the auditory feedback of speech, and using a closed-loop computer system to produce time-sampled feedback of eye-movement tracking and transformed feedback of hand motions. Closed-circuit television also has been used to study cybernetic control in performance and learning.

## Learning Feedback Control of Basic Movements

Recent advances in closed-loop computer techniques have made possible experimental analyses of changes in feedback control of primary body movements during the course of learning (K. U. Smith, Ansell, Koehler, & Servos, 1964a; K. U. Smith, Mysziewski, Mergen, & Koehler, 1963a). As indicated by the lower two diagrams in Fig. 2, a general-purpose high-speed scientific computer (Control Data Corporation, 160-A) is tied directly to behavioral and physiological feedback mechanisms by means of electronic transducing and converting components. For example, in studies of speech, the subject's speech sounds are transduced by a microphone, amplified, and converted to digital form by means of a variable analog-to-digital converter. The digitized electronic signal then is transmitted to the computer which has been programed to vary the feedback pattern in specified ways. This programed signal then is converted back into analog form by a digital-analog converter, filtered, amplified, and finally transmitted to the protected earphones of the subject, who hears his own speech as computer-controlled feedback.

In our cybernetic research laboratory, the digital computer and its analog converting systems are arranged so that they can receive transduced data from various experimental stations and control sensory feedback to the human or animal subjects located in these stations. In every case, the recorded data provide information about particular response patterns, such as eye movements, hand motions, or breath control, and thus can be used by the subject as feedback to control his response. However, these feedback signals are presented to the subject only after they have been manipulated by the computer. Thus, each of the computer-controlled systems is a closed-loop operation in

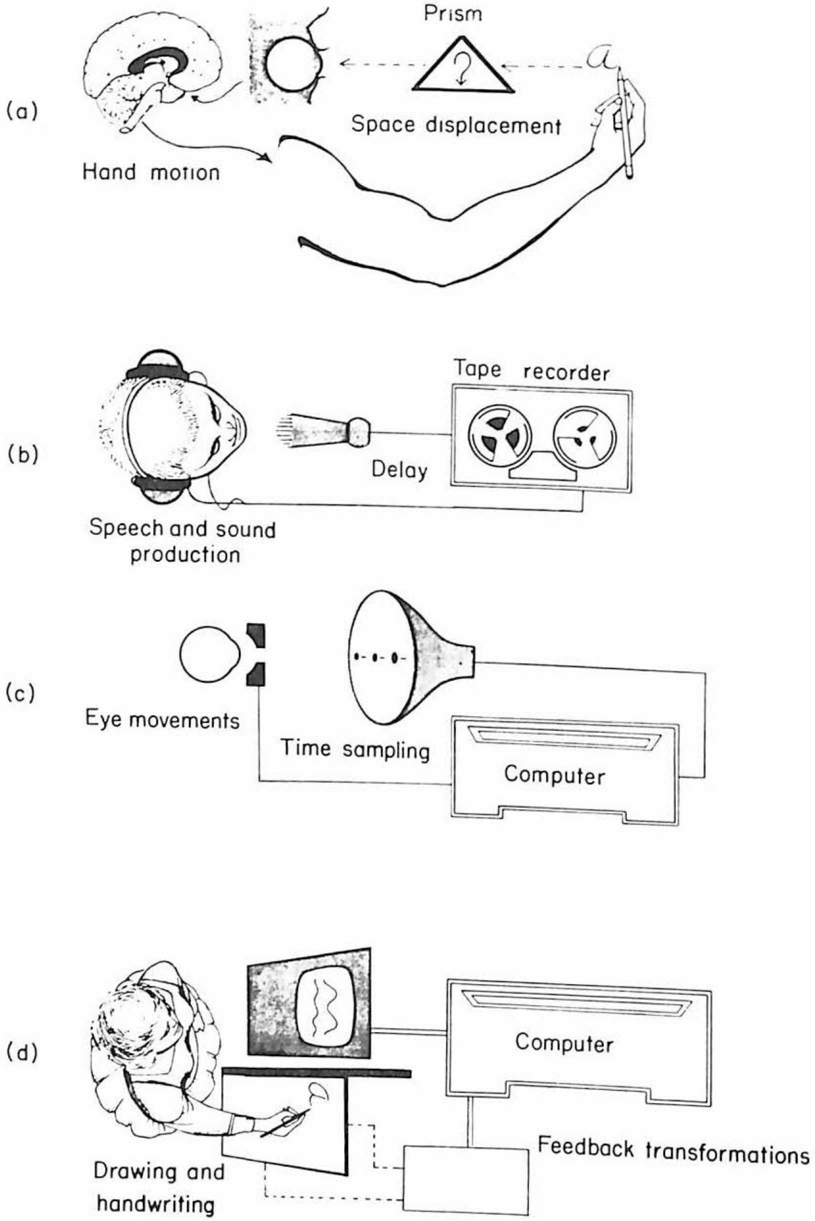


FIG. 2. Methods of experimental cybernetic research. (a) Optical system for controlling spatial displacement of the visual feedback of motion. (b) Magnetic tape system for delaying the auditory feedback of speech. (c) Variable control of time-sampled feedback of eye-movement tracking. (d) Transformation of the feedback of hand motion by means of closed-loop computer methods.

which the subject regulates his own movements by means of a perceptual display that presents experimentally varied feedback patterns. Some of the variations that can be introduced include feedback delays, time-sampling, spatial inversions, temporal inversions, kinetic variations, signal distortions, signal transformations, and symbolic transformations. The effects of variations or perturbations of this sort are measured to obtain data indicating the functions and limits of cybernetic control of particular behavioral or physiological mechanisms. The computer is used not only to manipulate feedback signals, but also for various procedures of experimental monitoring, calibration, experimental measurement, on-line data analysis, and data summarization.

## LEARNING FEEDBACK CONTROL OF BREATH MOVEMENTS

The experimental setup diagramed in Fig. 3 shows how the closed-loop computer system is used to study a basic control pattern, that of maintaining a constant intraoral breath pressure, such as is used in speech, singing, and certain types of musical instrumentation. The subject attempts to maintain a

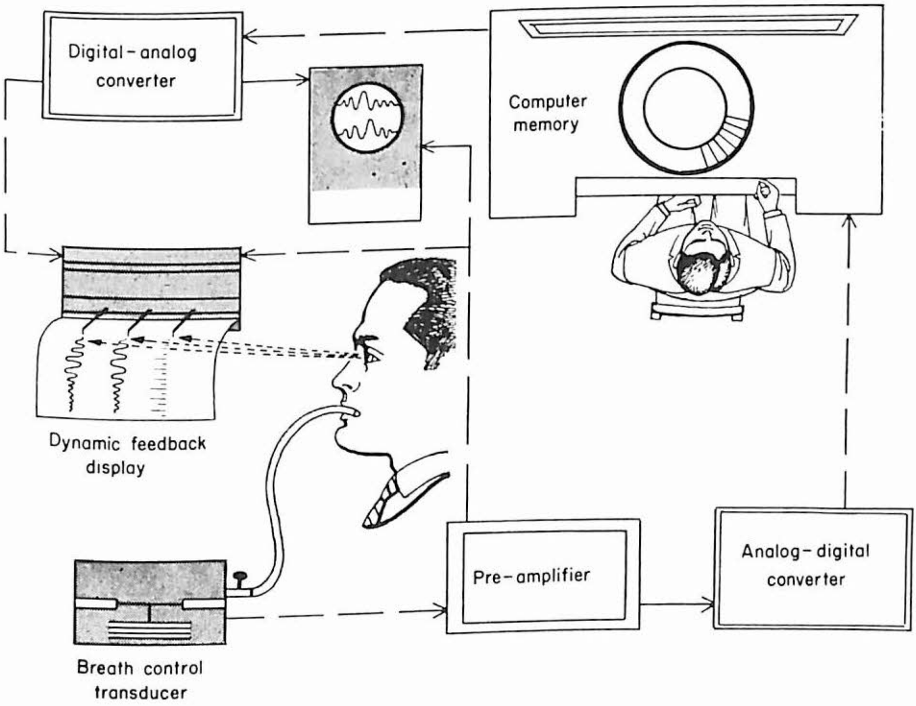


FIG. 3. Experimental cybernetic setup for studying learning of breath-pressure control with variable feedback.

constant breath pressure against a closed metal tambour which is fitted with a strain-gauge transducer. The transduced pressure signal is amplified, converted by the analog-digital converter, programed for quantitative variations in feedback, deconverted to its original analog form, and displayed on an oscillograph as a computer-controlled visual feedback display. The subject tries to make this oscilloscopic display of breath-control activity appear as a straight line.

Needless to say, the subject never achieves perfectly stable control over his response output and its related feedback input. The typical breath-control record in Fig. 4 is a very unstable pattern, marked by tremor and many other reactions of the sort that usually are called "spontaneous." Successive records show that as the subject learns better control of the activity, the magnitude of the oscillatory activity is reduced, the velocity of the component frequencies is reduced, and some of the larger oscillations disappear altogether. Tremor and oscillatory activity are characteristic of all movement systems of the body, with tremor rate varying in different response systems.

## COMPONENT MOVEMENT ANALYSES OF CYBERNETIC CONTROL

Analysis of the frequency characteristics of the breath-pressure pattern discloses that it is made up of a number of component activities. Further analysis shows that some of the variations in the record of somatic movement are perturbations introduced by internal physiological activities and that learning feedback control of the external responses involves modifications of the physiological perturbations.

One way to analyze the activity components is to use the computer system to differentiate the primary movement pattern and thus to determine the velocity properties of the responses. Velocity records of breath-control movements show a rhythmic variation which corresponds to the heart rhythm and larger variations which are introduced by the respiratory cycle. These physiological perturbations consist of both ballistocardiographic and ballistopneumographic variations in muscle action as well as neurally determined perturbations of the same origins, and represent an unavoidable interaction between internally organized activities and somatic sensorimotor control patterns. Velocity records obtained over a series of trials indicate that learning greatly reduces the range of velocities.

Additional observations have shown that some of the variations in the breath-pressure control pattern are related to the different component movement systems of the lips, tongue, thorax, and abdomen. The nature of the breath-control pattern shifts as movements of these different parts of the body are used to regulate the breath pressure. According to our interpretation, the oscillations due to these different activities reveal the differential role played by articulated manipulative reactions and background transport and postural activities in maintaining breath control.

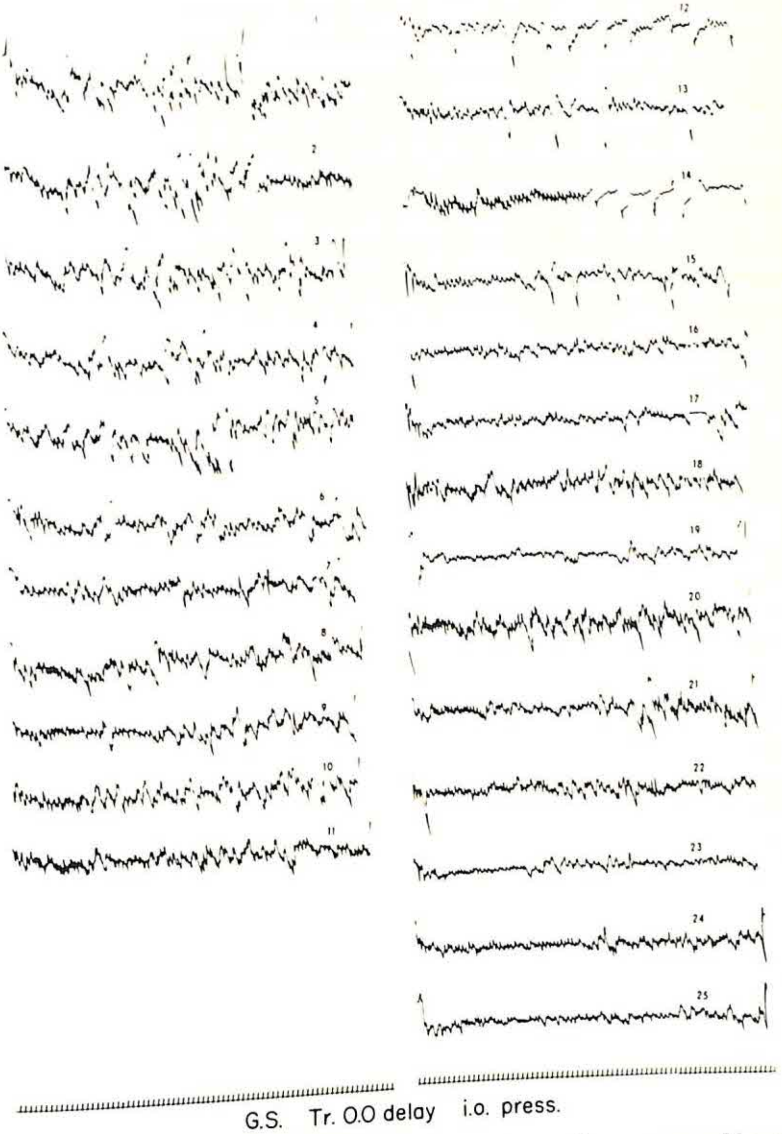


FIG. 4. Changes in breath-control movements during the process of learning as recorded on-line by the oscillograph of the closed-loop computer system. The strips of record have been cut from the original and mounted in order of trials. Time is in seconds. (Experiment in collaboration with Dr. J. P. Henry & R. Junas.)

Research of this sort gives a picture of the changes involved in somatic learning which is altogether different from the accounts produced by traditional learning studies. These analyses of feedback control reveal that all somatic responses involve continually generated tremor activities due to either the ballistic or neural actions of the cardiac and respiratory systems as well as grosser perturbations of response due to these systems. Thus learning a pattern of exteroceptive feedback control involves integrating somatic movements with these background activities and perturbations. Furthermore, the somatic movements themselves are made up of different component actions, including the relatively slow movement frequencies related to the postural and transport systems and the faster and smaller manipulative or positioning reactions. The whole motion pattern is integrated systematically by sensory-feedback mechanisms which determine the relative prominence of different movements.

In learning to integrate the internal physiological perturbations with somatic movements to achieve a more effective and accurate pattern of response, the individual utilizes the immediate feedback effects to modify both internal and somatic behavior. In this picture of learning change, fortuitous temporal contiguities and extrinsic reinforcements play no major role.

## Delayed Feedback and Learning

Delayed feedback refers to the introduction of temporal lags between a movement and some source of sensory feedback giving information about the effects of the movement (K. U. Smith, 1962). This can be done experimentally by contriving to interrupt the normal feedback loop. Delayed feedback is a feature of some kinds of machine performance, where the design of the machine introduces a lag between the operator's movements and their perceived effects, and also characterizes certain remote-control situations, such as the operation of a remote space vehicle from a station on earth.

The problem of delayed feedback is a crucial one for learning theory because it deals with the basic issue of whether or not plasticity of behavior in learning can be attributed to the presumed ability of the behaving organism to associate or link discrete events that are related only by temporal contiguity. Both contiguity theory and reinforcement theory, which in their several variations represent the only significant S-R learning theories currently in vogue, are based on the old assumption of associationism that events (stimuli or responses) occurring close together in time can be "associated" in behavior on the basis of the temporal linkage alone. The special condition imposed by reinforcement theory—that a reinforcement must occur to ensure the association—does not change the original assumption of temporally determined shifts in behavior patterning, for the only relation the reinforcement need bear to the response is that of temporal contiguity.

As the various S-R expressions of associationism have taken form in learning theory and research, the temporal factor common to all of them has not been defined or measured precisely. Temporal contiguity in classical conditioning means an interval (between conditioned stimulus and unconditioned stimulus) of variable length, often about .5 second. In operant conditioning, an interval (between R and reinforcer) of zero is assumed to be optimal but it can be stretched to seconds or in some cases even to minutes without destroying its effectiveness as a determinant of learning. In other words, conventional open-loop learning models of necessity assume a high degree of flexibility in the temporal relationships that constitute "contiguity" and thus in these theories determine the learning process.

Investigations of the effects of delayed sensory feedback on learning offer a serious challenge both to the open-loop conception of the learning interaction and to the notion that variable time intervals are the definitive factor in behavior reorganization. The delayed feedback experiment not only attests to the closed-loop nature of response but also indicates that the primary stimulus-response interaction has very rigidly defined temporal relationships. Research on delayed feedback does not eliminate contiguity or reinforcement as possible factors in reorganizing behavior, but it offers convincing evidence that fortuitous temporal relationships or open-loop contingencies do not *of themselves* determine response patterning or learning change.

Studies of transmission lags in velocity and aided tracking systems during and after World War II indicated that any lag between movement and perceived effect is detrimental to performance accuracy (Conklin, 1957; Levine, 1953; Lincoln & Smith, 1951; Warrick, 1949). Lee (1950a, b, 1951) originated another type of delayed feedback experiment by using audio-tape systems to delay the auditory feedback of speech and other sound-producing motions. In this technique, the subject is fitted with protective earphones in order to exclude all airborne sounds except those conveyed by a closed-loop audio-tape system. The magnetic tape stores the transduced speech and plays it back as a delayed feedback to the ears of the speaker. Lee reported that delayed side-tone induced stuttering, marked slowing of speech, and sometimes complete blocking. Other sound-producing motions were affected similarly.

A long series of studies on delayed auditory feedback has confirmed Lee's general results and has established many more specific findings concerning the nature of the induced disturbance (K. U. Smith, 1962; Yates, 1963). A few studies which investigated the nature of learning under the delay condition showed that limited improvement in performance can occur (Tiffany & Hanley, 1956; Winchester, Gibbons, & Krebs, 1959). Artificial stutter induced by the delay is reduced, and the slowing in speech which occurs initially may become less marked with practice. However, persons who learn to overcome the effects of the delay apparently do so by ignoring the delayed sound and using a different channel of feedback control (Goldiamond, Atkinson, & Bilger, 1962).

## COMPUTER-CONTROLLED DELAYED FEEDBACK

A computer-controlled closed-loop system has several significant advantages in delayed-feedback research. Besides providing precise control of delay intervals over a considerable range, such a system makes it possible to study both visual and auditory delay and to compare the effects of delay with the effects of other feedback variations (Smith *et al.*, 1964a).

A computer system used in delayed-feedback studies is diagramed in Fig. 5. For speech research, the subject sits in a small sound-resistant room, speaks into a precision microphone, and receives auditory feedback through protected earphones. The speech is recorded on a tape recorder located outside the subject's room and also may be displayed on an oscilloscope visible to the experimenter. To condition the microphone signal for computer programing, it is amplified, filtered if necessary, and then put through a small analog computer which samples the speech signal at rates up to approximately 10,000 cycles per second and converts it to digital form. This digitized signal is delayed by the computer, then deconverted to analog form, filtered, amplified, and presented to the subject as delayed auditory feedback.

Programing the computer to control feedback delay can be conceptualized as a process of electronically writing in and reading out the digitized feedback signal on the localized spaces of the iron memory core. The delay is produced by spatially separating the read-out point from the point at which write-in of the signal occurred, and the length of the delay is determined by the distance between the two points. Minimum delay (the feedback condition of "no delay") is obtained by reading out a signal at the same core point on which a write-in has just been made. The minimum delay in our present setup is .0002 second. Maximum delay is obtained by writing in at a point just beyond that at which reading out occurred in the previous programing cycle. The actual maximum thus depends on the storage capacity of the memory core, which in this case is 19,880 bits, and the rate at which bits of information are written in. Using a sampling frequency of about 10,000 samples per second, we secure a maximum delay of about 2 seconds. Longer delays can be obtained by reducing the sampling rate. We have been especially in-

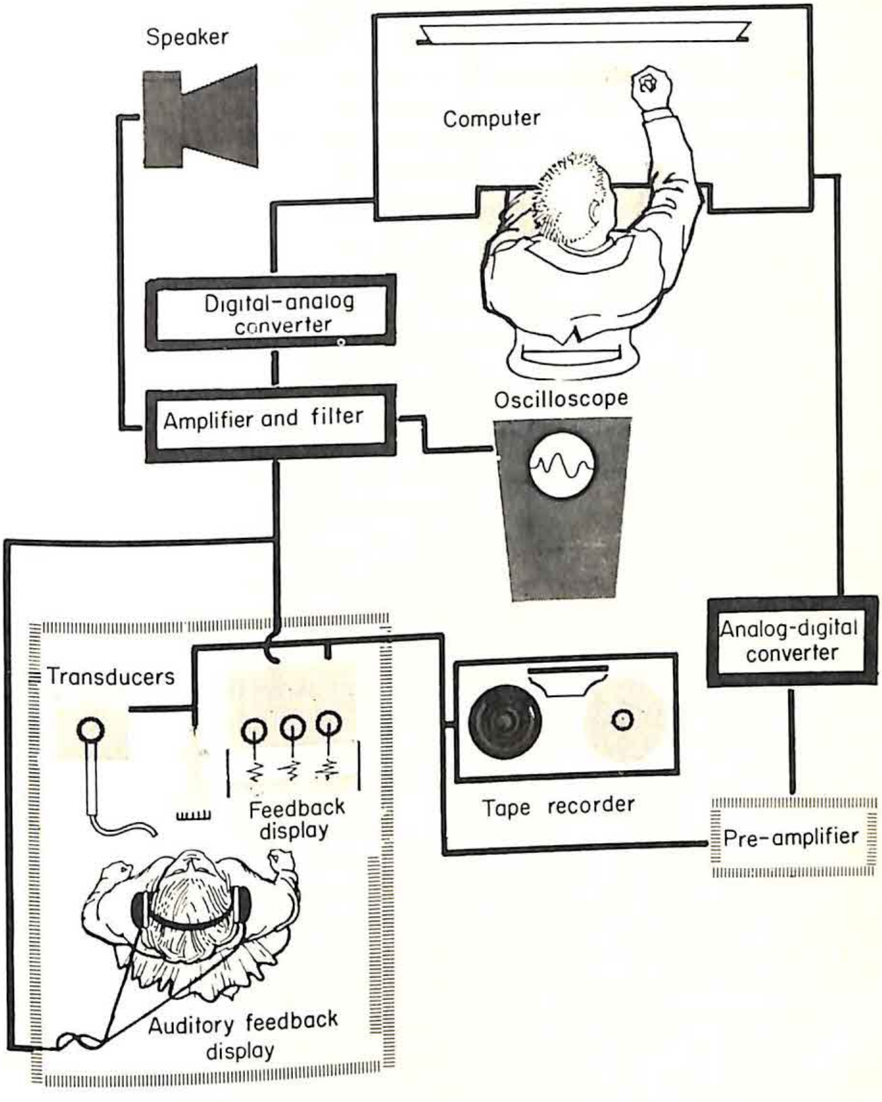


FIG. 5. Using a computer system for research on delayed auditory feedback. Electronic converters and a general-purpose scientific computer (Control Data Corporation 160-A) are used to convert, sample, preprogram, and deconvert the sound signal in a closed-circuit auditory feedback loop. The sounds of speech and other forms of sound production can be delayed or otherwise perturbed before being presented as auditory feedback. (Based on K. U. Smith *et al.*, 1964a.)

terested in feedback delays from about 1.2 to 2.7 seconds because any communication between earth-bound control stations and space vehicles on the moon would be delayed by intervals within this range. The time needed to transmit signals to the moon and to receive returning signals ranges from 2.38 to 2.71 seconds, depending on orbital positions.

Our present computer system is designed to perform various kinds of experimental measurements and to control different read-out checks on the speech signal in addition to delaying or otherwise manipulating the signal. Oscilloscope and oscillograph read-out checks can be made at selected points in the analog-digital and digital-analog converters, which make it possible to determine at any time the accuracy of input-output control of the system.

### Learning with Computer-Delayed Auditory Feedback

A number of experiments have been carried out to compare the effects of delay with the effects of other variations of the feedback signal. The effects of computer-controlled delay also have been compared with effects obtained with audio-tape systems. Results on delayed speech were obtained from 10 subjects who were either graduate students or members of the laboratory group (including the author of this chapter). In all observations, subjects were asked to read or speak on fairly technical subjects.

To our knowledge, this was the first successful use of the digital computer to transmit human speech with fidelity, and the quality of the signal exceeded our expectations. The system transmits the speaker's voice with its individual voice quality and retains subtle features of enunciation. Two voices transmitted over the system simultaneously are heard as two distinct voices. Computer-controlled speech seems to be somewhat lower in pitch than the normal voice and also has a somewhat more resonant quality, but proved entirely satisfactory for our purposes. Instrumental music also can be controlled by the computer with a high degree of fidelity.

In general, computer-delayed speech feedback produces effects similar to those produced with the audio-tape technique. In some cases, the computer technique seems to produce more striking results than audio tape, possibly because the computer transmits speech sounds with higher fidelity. Both techniques produce a peak disturbance with delay magnitudes around .2 second.

There are marked individual differences in response to delayed auditory feedback, and the severity of the effects appears to increase with age. Individual patterns vary from slight slowing of speech and increase in intensity to radically disorganized speech that resembles that found with some neurological disorders. Stuttering, blocking, emotional voice quality, and other radical alterations in the speech pattern are observed. Individuals who show severe disturbances are affected by a wider range of delay intervals than are those persons who display only minor effects.

Repeated exposure to the delayed speech produces many modes of adjustment in different subjects. Some persons improve their speech with practice while others show little or no change and still others may get worse. Although many persons can improve if required to read only short sentences, continuous reading may produce poorer and poorer performance. Some persons who show only limited initial reaction to the delay may subsequently break down and then be unable to reestablish control. Persons who can perform under the delay condition with few speech disturbances seem to find the task even more demanding than individuals who show marked disturbance. This suggests that overcoming the effects of the delay may involve other channels of feedback control which are more difficult to monitor than the auditory channel. There appears to be no true learning to react to the delayed signal, and any adaptation that occurs apparently involves establishing some other pattern of feedback control.

We have found that material read under delayed feedback conditions is retained better by subjects who show marked disturbances than by those who are affected only slightly. This result seems to indicate that persons who are able to maintain their speech are controlling it by sensory-feedback signals which do not convey systematic meanings.

We have used the computer system to study delayed auditory feedback of a number of sound-producing responses other than speech. Delayed whistling and singing produce blocks and slowing and slurring of sound production. The effects of delaying the auditory feedback of musical instrumentation vary with the type of instrument. More marked effects are noted with instruments requiring continuous control of sound, such as a wind instrument, than with instruments such as the piano on which discrete sounds are produced in series. In some cases, a peak disturbance can be found with delay intervals of about .4 second or above, in contrast to the peaking effect at .2 second for delayed speech. However, with some instruments the severity of disturbance increases as a function of the magnitude of the delay.

Studies of musical instrumentation with delayed feedback are of particular interest because results can be obtained from persons who have achieved varying levels of proficiency in performance. Expert musicians are affected far more by delayed auditory feedback than novices and refuse to perform repeatedly under the delay condition. Also, more marked effects are found with difficult musical scores than with simple scores. These results indicate that the severity of the disturbance may vary directly with the precision of control normally required in the activity. There are reasons to believe that this relationship also holds for delayed speech. Subjects (musical novices) who practice piano playing with delayed feedback show only limited improvement.

We have carried out learning experiments with delayed auditory feedback by setting up tasks such as tracking or tracing in which errors produce sound signals. Results indicate that adaptation to the delayed signals is largely a matter of finding some secondary channel of feedback control which can be used to regulate response and to promote learning. For example, the availability of proprioceptive and cutaneous feedback in tracking tasks makes it possible to learn even when the auditory error signal is delayed.

The computer system makes possible direct comparisons between delay and other experimental manipulations of the auditory feedback signal. Programs can be written to interrupt the sound at various intervals, to produce artificial repetitions in speech, to invert specified intervals of speech, and to distort the feedback signal in various ways. Preliminary study of some of these feedback variations suggests that the informational or time-sampling interpretation of speech control, which attempts to describe the perceptual nature of speech in terms of stimulus properties which can be specified as binary bits of information, misses the important characteristics of speech as an organized closed-loop response system. The information model makes no provision for the temporal requirements of the speech transmission system and thus cannot predict the disturbance effected by delayed auditory feedback or provide hypotheses to be tested in this area of research. The phenomena of delayed auditory feedback are just as foreign to the formulations of information theory as they are to conventional association, conditioning, and reinforcement doctrine.

## Learning Breath Control with Delayed Visual Feedback

In a previous section, we described how we use the computer system to control a visual feedback display of breath control. Using this technique, we delayed the visual feedback for various intervals and obtained records of breath control over a series of practice trials. Figure 6 presents a series of 1-minute records with visual feedback delays between .0 and 3.2 seconds obtained from a subject fairly late in a practice series. It will be observed that the basic pattern of the movements changes considerably as a function of the delay magnitude. As the delay is increased up to 1.5 and 2.0 seconds, the magnitude of error increases. Extracting the first derivative or velocity component of these records reveals that changes in velocity follow a different function. With increased magnitude of feedback delay, the velocity of the control movements becomes more variable and also decreases in magnitude.

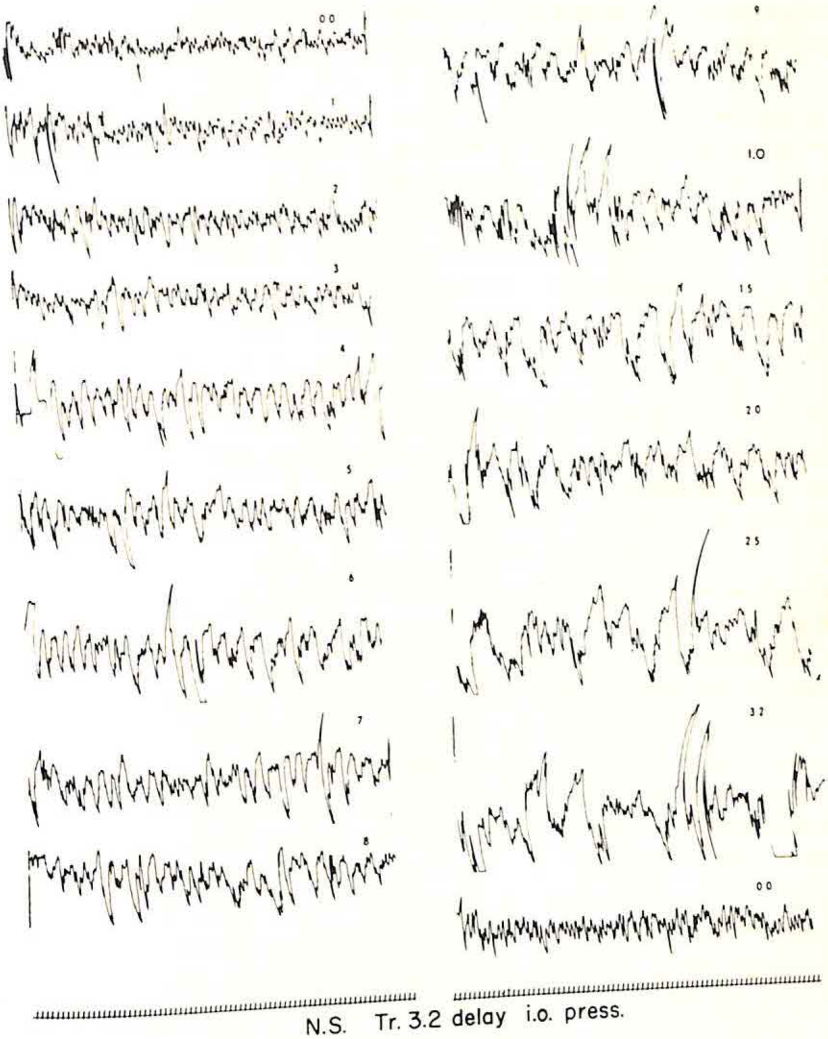


FIG. 6. Variation in the pattern of movement with different magnitudes of delayed feedback. (Experiment in collaboration with Dr. J. P. Henry & R. Junas.)

The records indicate that the pattern of neurobehavioral control varies systematically with increases in magnitude of feedback delay.

The nature of learning change under conditions of delayed feedback was studied by obtaining practice series of breath-control records with delayed visual feedback. A comparable series of records obtained with zero delay was shown in Fig. 4. Improvement occurred under both conditions, but the final level of accuracy achieved with feedback delay did not equal the level of control learned with no delay. Some positive transfer of learning was found from one delay interval to others.

Both performance and learning under the delay condition were highly unstable and subject to irregular variations and regressions to earlier levels. Transfer effects also were highly variable and unpredictable. The whole picture of response and learning with delayed feedback is that of an unstable system marked by irregularities of control and periodic lapses of control. These effects are as marked with delayed visual feedback as with delayed auditory feedback.

### Learning with Delayed Televised Feedback

A limited number of delayed feedback studies have been carried out with closed-circuit television techniques in order to delay a true pictorial feedback image (K. U. Smith, 1962). One study involved presenting a simulated feedback image to the subject in the manner shown in Fig. 7. Two closed-circuit camera-monitor chains were used, one of which transmitted a televised image of the subject's manual performance to the monitor of a trained observer, and the other of which transmitted an image of the observer's performance to the subject's monitor. The subject was told that he was seeing an image of his own hand and arm delayed electronically. He was instructed to trace a maze by watching the hand performing in the monitor image. Actually, what he saw was the image of the observer's hand, duplicating the subject's performance. White gloves and black arm covers worn by both subject and observer aided in the deception.

With this simulated feedback technique, it was possible to delay the feedback image watched by the subject for different intervals. The observer started his delayed performance either after a short interval of about 1 to 2 seconds, after the subject was half-way through the maze-tracing task (about 15 seconds), or after the subject had completed the task (about 30 seconds). Four subjects practiced maze tracing at each of these delay conditions for five trials per day for three days. Two types of accuracy scores for the three groups as a function of practice periods are shown in Fig. 7. The very slight improvement shown was not significant.

A number of experiments were carried out using a video-tape recorder to record the televised image of performance and then to play it back to the subject's monitor after a task had been completed. Subjects were asked to practice tracing and dotting tasks "blind," and then were shown the delayed recorded feedback image at the conclusion of each trial. Only the most limited learning occurred under this condition.

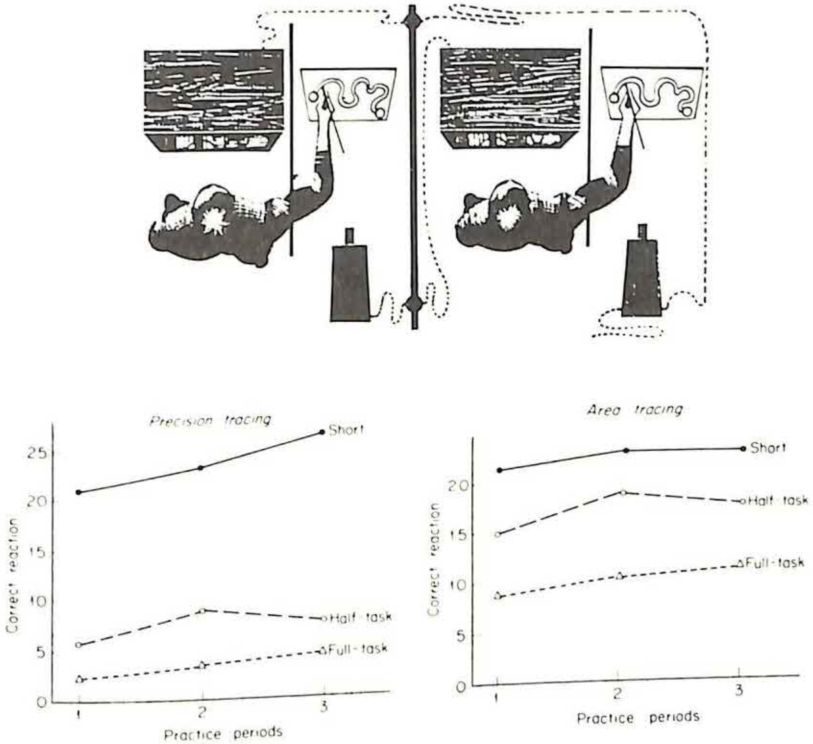


FIG. 7. Method and results of a study of learning with variable delayed feedback simulated by means of a dual camera-monitor television chain. The curves show the two types of accuracy scores in a simple maze-tracing task as a function of practice with three different magnitudes of concurrent delayed feedback (approximately 1, 15, and 30 seconds). (Based on K. U. Smith & Smith, 1962.)

## FEEDBACK DELAY AND REINFORCEMENT DELAY IN LEARNING

The significance of the feedback concept is recognized widely, but there is an unfortunate tendency among psychologists to equate feedback with knowledge of results and also with reinforcement. Delayed feedback then is assumed to be equivalent to delayed reinforcement, and individuals are said to show marked disturbances of performance and learning with delayed feedback because their speech, handwriting, and other movements were learned originally with reinforcement delays different from those used in the feedback experiments.

The possibly erroneous nature of these assumptions can be shown by comparing the empirically determined effects of reinforcement delay and feedback delay. It should be noted that it is very difficult to demonstrate detrimental effects of reinforcement delay in human somatic behavior, and no systematic functions for human subjects have been published. However, the delayed reinforcement data obtained by Wolfe (1934) on rats were plotted by Hull (1943) to show decreasing effectiveness of reinforcement as a function of delay interval. Although response strength dropped off rapidly with delays between 0 and 30 seconds, the decline was much more gradual with longer delays. With a 1-minute delay, response strength was plotted at a level one-third as high as maximal, and some learning occurred even when reinforcement was delayed as long as 20 minutes. The results of delayed feedback experiments show nothing comparable with these results on delayed reinforcement. Most of the major disturbances of performance and learning with delayed feedback occur with delay intervals of less than 1 second. Furthermore, the disturbances of performance control that occur with delayed feedback do not appear with delayed reinforcement. The obvious reason for this difference is that reinforcing events are extrinsic to the sensorimotor mechanism whereas feedback is an intrinsic component of response, necessary for its control.

In our opinion, the major conclusion that should be drawn from the research on delayed feedback is that disturbances induced in behavior and learning are the result of perturbation of the immediate movement-produced sensory feedback stimuli. These stimuli are absolutely essential for closed-loop response control and learning and are not equivalent to reinforcement or to static after-effects providing knowledge of results. Explaining the effects of delayed feedback as due to a disturbance of learned open-loop contingencies rests on a number of gratuitous assumptions and discloses an ignorance of the nature of closed-loop cybernetic control in behavior. Any pattern of movement, whether of the receptor or of the somatic motor system, is controlled not by "perception" or by motor factors alone, but by continual perceptual-motor interactions provided by the feedback mechanisms. When any property of the controlling feedback pattern is altered, as in temporal delay, spatial displacement, kinetic transformation, or time sampling, the form of the response is changed. One of the effects of learning is to modulate the pattern of response—for example, by introducing compensatory movements, or by changing from a continuous pattern to discrete movements—in order to adapt the control pattern to perturbed feedback conditions. However, research results show that only the most limited adaptation can be made to the temporal perturbation of delay.

## Cybernetic Analysis of the Spatial Factors in Learning

Since the time of Helmholtz (1925), it has been recognized widely that one of the most prominent features of behavior in perception and learning is its spatial organization. Yet outstanding behavior theorists, including Helmholtz, have tried to explain spatial organization as a by-product of temporal organization. All associative learning theories as well as the accepted doctrine of synaptic integration of response assume that the behaving system differentiates and integrates events on the basis of their temporal relationships—i.e., that the system somehow is sensitive to time itself. Such a view poses enormous logical problems that to our mind are chimerical.

One of the main contributions of the cybernetic approach is to call attention to the primary spatial organization of response, to devise methods for analyzing it, and to suggest neural mechanisms whereby it might logically be mediated (K. U. Smith, 1961a, b; K. U. Smith & Smith, 1962). A first important step was to recognize that the classical studies of displaced vision (Brown, 1928; Ewert, 1930; Peterson & Peterson, 1938; Siipola, 1935; Stratton, 1896, 1897, 1899; Wooster, 1923) actually originated the area of cybernetic research. These studies of geometric displacement of movement-generated stimuli were extended by better controlled optical methods (Rhule & Smith, 1959), more flexible television techniques (K. U. Smith, 1961b; K. U. Smith & Smith, 1962; W. M. Smith, Smith, Stanley, & Harley, 1956), and closed-loop computer methods (K. U. Smith, *et al.*, 1964a; K. U. Smith & Smith, 1966).

### LEARNING WITH ANGULAR DISPLACEMENT OF FEEDBACK

Cybernetic theory implies directional guidance of response based on detection of differences in stimulus patterns. Behavior control can be likened to control by electromechanical servosystems, which orient a movable component according to detected differences in input patterns. However, control of a behaving system is greatly complicated by the complexity of the system. It has many input channels as well as multidimensional patterns of response. Furthermore, its mobile receptors and other moving parts constantly change position with relation to each other, thus spatially displacing the stimulus feedback patterns by means of which control is maintained. However, the behaving system can maintain normal efficiency and accuracy of performance within what we call the "normal" angular range of displacement, the actual extent of which varies with the receptor system, the kind of response pattern, its precision, its complexity, and other variables. Angular displacement beyond this normal range causes breakdown in perceptual orientation, pattern recognition, motor control, and learning, the degree of which will vary as a function of the magnitude of the feedback displacement. The concepts of normal and breakdown ranges of angular displacement were introduced earlier and diagramed in Fig. 1. One of the important functions of learning is to extend the normal displacement range—that is, the range within which performance efficiency can be maintained.

These ideas constitute a central hypothesis of our version of cybernetic theory and have been tested in controlled studies of performance in displaced televised fields. A clear demonstration of the breakdown function was obtained in an experiment on maze tracing illustrated in Fig. 8. Subjects were trained for 9 days to trace a paper-and-pencil maze by watching their performance in a television monitor. Four different camera placements during training displaced the locus of vision 0, 20, 40, and 60 degrees from normal. On the 10th day, the subjects were tested for accuracy with the eight camera placements shown in the figure. The curve in the figure shows the mean number of errors (in tracing cleanly through the short passageways of the maze) plotted as a function of visual displacement. It can be seen that error level remained constant with visual displacements through 40 degrees but then rose sharply with larger displacements. In experiments of this sort, the extent of the so-called normal range of displacement depends in part on the difficulty of the task and the precision required. Further, the efficiency of performance with different spatial displacements has been shown to improve with practice (K. U. Smith & Smith, 1962; K. U. Smith, Wargo, Jones, & Smith, 1963b). In one experiment (McDermid & Smith, 1964), learning functions were obtained for dotting, drawing, and handwriting motions performed either directly (normally oriented) or in compensatory fashion (oriented so as to look normal) with televised feedback images displaced 0, 90, 180, and 270 degrees from the normal locus. Two main assumptions were being tested, namely: (1) that learning functions will vary with both magnitude of displacement angle and mode of response; (2) that compensatory response or negative feedback control will be less accurate than direct response inasmuch as compensatory performance involves more feedback discrepancies than the direct mode of reaction.

The results of this experiment, as tested by Duncan's (1955) multiple range tests, showed significant variations in movement times for both mode of response and magnitude of displacement. Level of error and degree of learning change were greatest for compensatory response at 90- and 180-degree displacement angles. With the direct mode of response, the variation due to displacement angle was not great except for the difference between the 0-degree displacement and the other displacement conditions. The 270-degree displacement condition (equivalent to 90 degrees to the right) caused little difficulty, probably because right-handed writing normally slants to the right and thus is controlled easily by right-handers with the 270-degree feedback angle. Performance varied significantly for tasks of different complexity (there were no significant differences for the dotting task), for different movement components (travel and manipulation), and as a function of practice. With the large displacement conditions used in this experiment, draw-

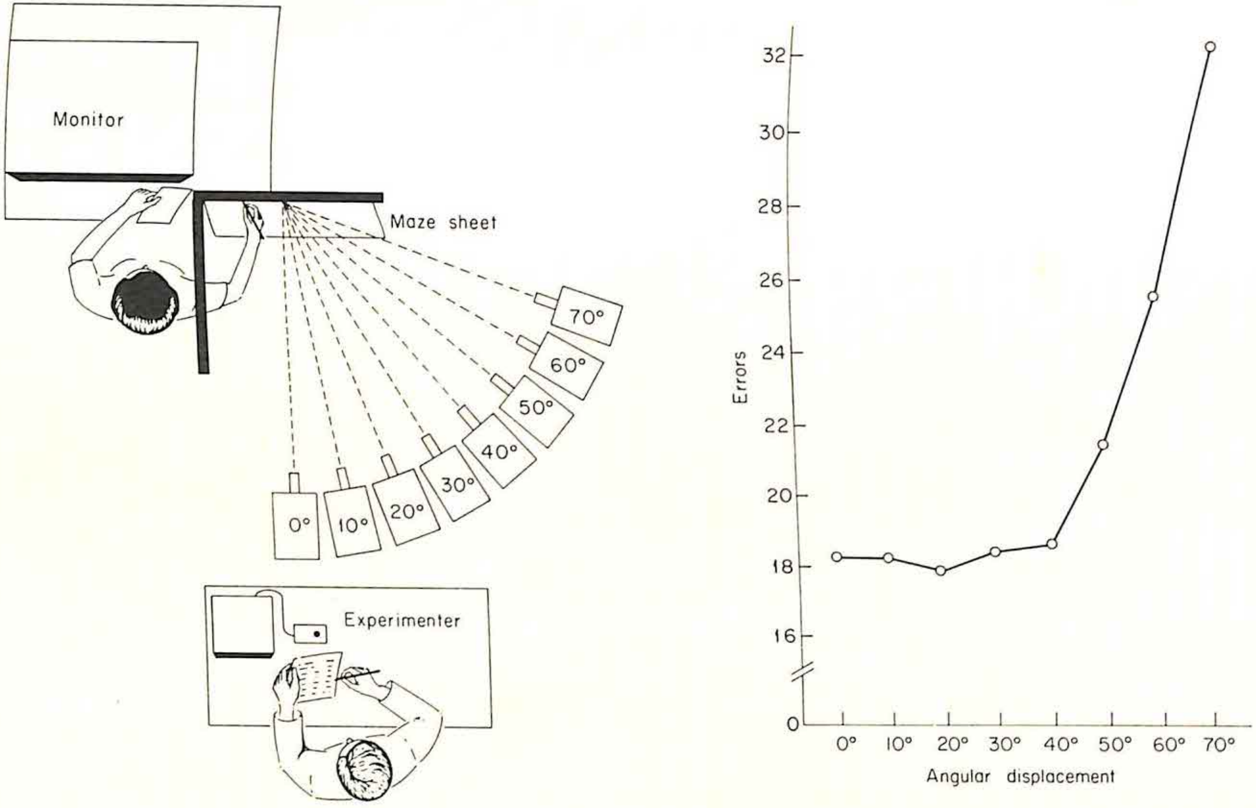


FIG. 8. Method and results of study demonstrating normal and breakdown ranges of control in maze tracing as a function of angular displacement of visual feedback. (From Gould & Smith, 1963.)

ing and writing times were still elevated above normal after 9 days of practice but the mean differences had been reduced considerably. The behaving system can adapt fairly well to spatial displacements of feedback patterns although, as we have seen, there is little or no adaptation to temporal delay.

The data of this experiment showed that practice increased the number of significant differences in performance times even though the size of the differences was reduced. This indicates that responses became more sharply differentiated relative to directional characteristics of the feedback-control patterns. Other observations have revealed similar direction-specific ranges of cybernetic control for particular receptor systems. For example, if we require a subject to read material located on a vertical surface before his eyes and then rotate the material to the right or to the left, different breakdown thresholds are obtained for the different directions of rotation. Similarly, one can determine breakdown thresholds for maintaining single vision of disparate stimuli by convergence and divergence movements of the two eyes.

These observations support the assumption of our neuronic theory that there are direction-specific neural detection mechanisms related both to the exteroceptive and proprioceptive systems of the body. Inasmuch as normal ranges of directional control may be large, it is likely that regional direction-specific neuron mechanisms of the brain could spread over a good part of a cortical projection area. Hubel (1963) and Hubel and Wiesel (1962) found in the striate areas of the cortex of cats direction-specific complexes of neurons which may represent mechanisms for visual control within certain feedback displacement ranges.

## Axial Coordinate Control Mechanisms

Another postulate of our cybernetic theory is that body motion is patterned in relation to environmental stimuli within a coordinate frame of reference provided by axially organized feedback-control systems. The idea is that the postural-gravitational and bilaterally symmetrical transport systems of the body constitute built-in vertical and lateral reference mechanisms to direct all movements with respect to the position of the body. Thus, the behaving system adjusts the direction of movements relative to the axes of the body while the system itself is changing position in space. One hypothesis to be tested in this regard is that displacements of feedback signals along the major axes will have differential effects on performance and learning. In a number of experiments testing the relative effects of visual inversions and reversals, we assumed further that inversions would effect greater disturbances than reversals because the postural movement system providing the up-down reference is more basic and phylogenetically older than the bilaterally referenced transport system.

Figure 9 illustrates an experiment on tracing movements, in which a subject controlled his own movements by watching a television monitor. The scanning circuits of the television camera were modified so that the television image of performance could be presented in normal, reversed, inverted, or inverted-reversed orientation. Since other factors remained the same while only these feedback conditions were varied, exact experimental comparisons could be made between the different patterns of closed-loop control. Subjects traced a star maze which had numerous grooved culs-de-sac along its edges to prevent easy tactual control and to provide an error score along with the time scores recorded by an electronic motion analyzer. Four groups of

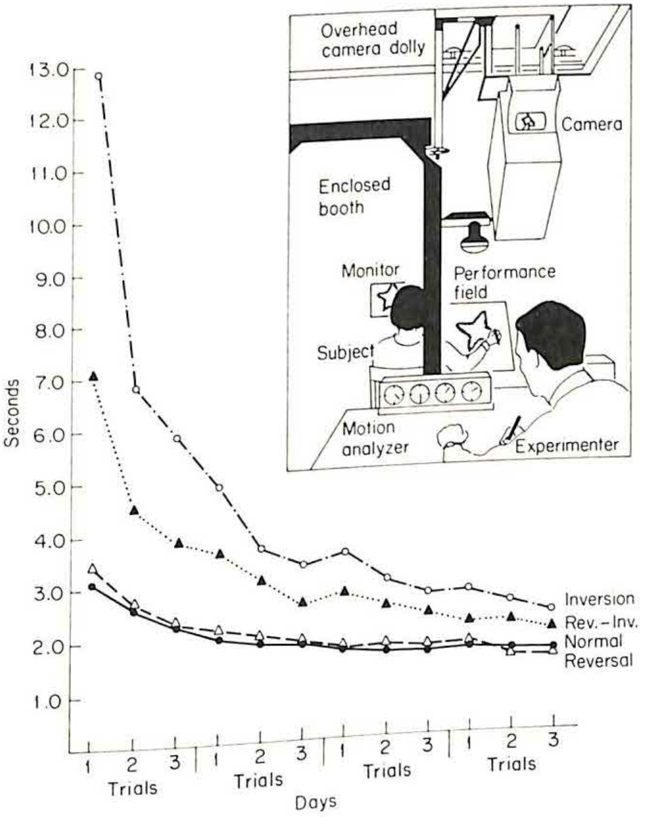


FIG. 9. Method and results of study of learning with visual feedback displacements in different axial dimensions. (From K. U. Smith & Smith, 1962.)

subjects practiced for 4 days, each group with a different orientation of the feedback images.

Learning curves for movement times are shown in Fig. 9. Each point plotted represents the mean for eight consecutive trials. Variance due to both practice and displacement conditions was significant. These data support our hypotheses that different displacement conditions will produce differential learning functions and that disturbance will be greater with inversion than with reversal. As expected, the inverted-reversed feedback did not impair performance and learning as much as inversion alone, presumably because the latter not only involves disturbance of the primary postural-gravitational reference system but also changes the normal coordinate relationships with the bilateral transport reference system. In this particular task, subjects performed as well with the right-left reversed feedback as they did with normally oriented feedback.

The differential effects noted in this experiment have been obtained in numerous similar studies on handwriting, tracing, drawing, target location, tracking, panel-control manipulations, gait, and other types of motion patterns. We believe that such results indicate decisively that regulation of motion involves coordinate control by separate feedback reference systems related to the up-down and right-left axes of the body. We also believe that these independently variable reference systems can be identified with the central postural system which regulates reactions in the vertical plane of action, and the bilaterally paired feedback mechanisms of transport movements. These interacting mechanisms would seem to provide a coordinate inertial reference system by means of which the space patterning of motion can be controlled in relation to specific external environmental patterns even as the body's position varies in space. In this view, learning psychomotor control involves integrating the feedback pattern of a particular manipulation made in relation to objects or external stimuli with the internalized reference feedback patterns of transport and postural motion. In order for a response to become efficient through learning, it must be controlled precisely by closed-loop processes along its spatial dimensions as well as in its temporal and kinetic characteristics. The neural implication of these ideas is that separate direction-specific regional detection systems of exteroceptive and proprioceptive control exist in the brain, and that these neuron detection systems are modified in learning by the feedback effects of movement.

## DIFFERENTIAL ADAPTATION TO DISPLACED FEEDBACK

A major postulate of cybernetic theory is that the relative complexity and precision of the motion pattern as well as the extent and dimension of feedback displacement affect the course of learning and the final level of performance in a task. In one experiment (Fig. 10), training in writing, drawing, and dotting tasks with prism-inverted visual feedback was extended over a period of 20 days. Twelve subjects performed either direct (normally oriented) or compensatory (written to look normal in the inverted field) re-

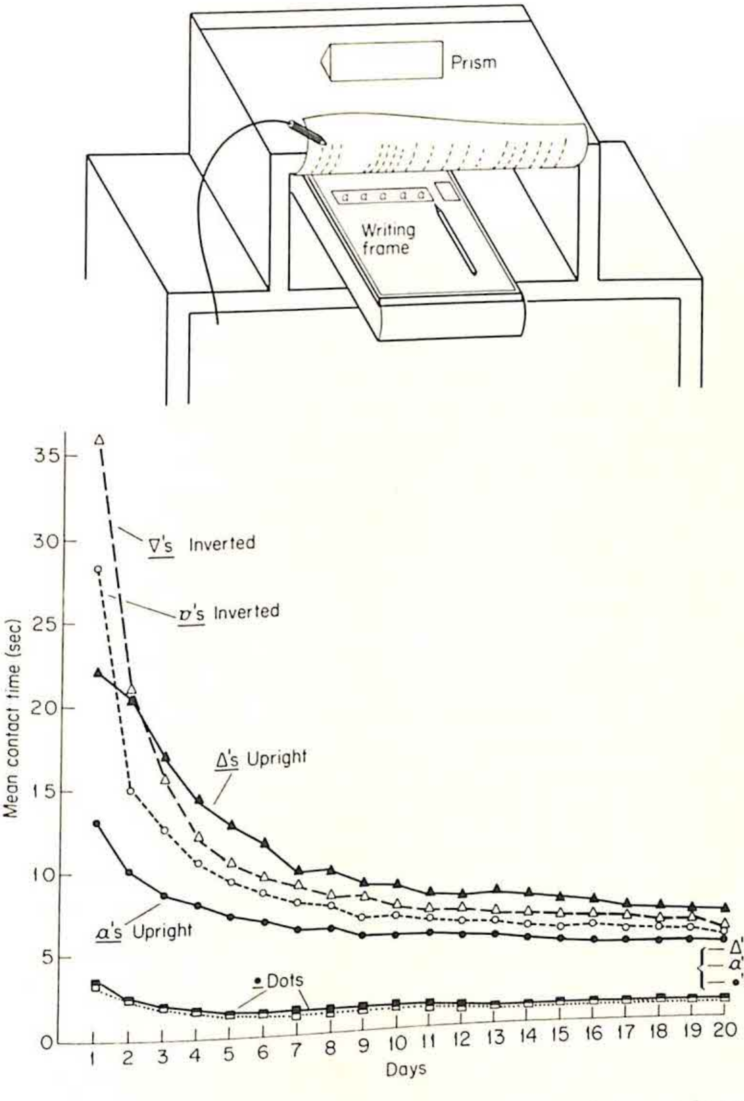


FIG. 10. Learning direct and compensatory responses with inverted visual feedback. (Based on K. U. Smith & Smith, 1962.)

sponses. Contact and travel movements were timed separately with an electronic motion analyzer. This experiment produced some of the most orderly learning functions which the writer has ever seen in motor-skill research. The curves in Fig. 10 show contact times for writing *a*'s, drawing triangles, and making series of dots in either direct or compensatory fashion. The course of learning varied significantly with the complexity of the task and with the type of response made in the inverted field. Movement times for writing and drawing did not reach the level established with normally oriented feedback even after 20 days of practice, and the differences between normal and displaced visual feedback were significant throughout the experiment.

These results challenge the assumption sometimes made that subjects given sufficient practice with inverting spectacles or prisms will show complete adaptation to the displaced vision (Kohler, 1951a, b, 1955; Snyder & Pronko, 1952; Held & Hein, 1958; Held & Schlank, 1959). As a matter of fact, careful examination of the data of prior experiments on displaced vision suggests that claims for generalized perceptual adaptation to inverted vision are unwarranted. These experiments have demonstrated limited adaptation of certain motion patterns, such as locomotion and object manipulation, but as Stratton's (1897) original studies indicated, adaptation never is complete.

All of the displaced feedback experiments suggest very strongly that plasticity of response in learning is based on spatial variations in the controlling feedback patterns. Cybernetic research investigates many parameters of feedback control – space displacement, intermittency, delay, kinetic variations, signal variations, and other systematic transformations. The delayed feedback studies have shown that responses are time specific, for the introduction of delay in the feedback loop degrades accuracy and prevents consistent learning. The only type of feedback variation that has been found to produce systematic variations in performance is the spatial one. Although experimental space displacements can produce marked initial disturbances in performance, the individual adapts to them at rates and to final levels that are specific functions of the dimensions and magnitudes of displacement. These statements apply to the learning of overt movement control as well as to receptor cybernetic control or perception.

## Spatial Factors in Response Specialization

An understanding of response specialization in learning is basic to our understanding of memory and transfer, for a response must become specialized or differentiated from other responses before it can be remembered or transferred. Cybernetic theory proposes that response specialization is based primarily on spatial properties of the controlling feedback patterns rather than the temporal factors of traditional learning theory. Several types of experiments which have been conducted to test this hypothesis all show that responses learned under one condition of space-displaced feedback are highly specific and transfer only in limited ways to other conditions of space-displaced feedback.

In one experiment, one group of subjects alternated practice of a response pattern with inverted feedback with practice under normal visual conditions. Their performance under the inversion condition did not differ from the performance of another group of subjects who practiced only with inverted feedback, indicating that there was no significant interaction in the first group between the inverted and normal visual training.

Similarly, transfer tests have shown that responses learned under one displacement condition show only limited transfer to other displacement conditions. When tracing, dot location, and drawing motions were practiced under inverted, reversed, or inverted-reversed visual feedback conditions, performance efficiency dropped markedly when the same responses were tested under the other feedback conditions.

Another experiment tested whether perceptual training under conditions of displaced visual feedback would transfer to related psychomotor tasks. Subjects practiced reading inverted material and then were tested for transfer of this training to inverted writing. Training in inverted reading had no discernible effect on the subjects' ability to write and draw with inverted vision—a finding that provides evidence of a high degree of specificity in perceptual learning as well as in overt psychomotor learning under particular spatial feedback conditions.

In a series of experiments subjects were trained in either direct (normally oriented) or compensatory response to various conditions of displaced vision and then tested in the other mode of reaction with the same displacement condition. Again, a high degree of spatial specialization was evidenced by the fact that little or no transfer occurred between the two response conditions.

Our most extensive analysis of directional specificity of feedback control compared the learning of movements in eight different directions under four different orientations of visual feedback. Subjects traced a four-pointed star with normal, reversed, inverted, or inverted-reversed visual feedback while a motion analyzer recorded movement times in eight different directions corresponding to the eight segments of the star. With four different feedback orientations and eight different directions of movement, it was possible to appraise the degree of functional differentiation in learning under 32 different spatial conditions of feedback control. Each group of eight subjects practiced under one of the four displacement conditions for 24 trials per

The cybernetic approach suggests a new basis for understanding similarity between two learning situations. If behavior is a closed-loop process, stimuli and responses should be considered not separate entities but interrelated events in a feedback loop. Similarity between activities and thus the degree of interaction between them depends not on isolated stimuli or isolated motor responses alone, but on relationships between patterns of feedback control. On the basis of present knowledge, we believe that the extent to which any specialized pattern of learned response will be utilized in some subsequent situation will be determined generally by the similarity of the space, time, kinetic, and transformational properties of feedback required in the two situations and the relative complexity and precision of the control patterns used in the two situations. We also believe that predicting specific transfer always will present difficulties because of the complexities involved in multidimensional response control and man's flexibility in shifting and modifying his pattern of control as a result of practice or of changed conditions. Because of the changes induced in the control of a learned task by practice, its similarity to another task also will vary. Inasmuch as skills involve postural, transport, and manipulative components, similarities and differences may exist at different levels of control. For example, two skills may be very similar in their pattern of transport movements and very different in their manipulative components; or similar in their manipulations and different in their transport pattern. Further, the wide individual differences that exist in control patterns of some tasks add greatly to the difficulties involved in predicting transfer effects.

The most general finding of all prior research on transfer of human learning is that generalization usually is very limited. This is especially true of the so-called psychomotor skills, tool using and machine operations. This specialization of learning reflects the ability of the individual to differentiate countless different spatial patterns of feedback control.

A transfer phenomenon that is more difficult to interpret is that of negative transfer, which indicates not just difference but interference between two response patterns. Negative transfer occurs not between extremely dissimilar or unrelated tasks, but when some elements are similar and some dissimilar. For example, transfer may be negative when the transfer task requires the same general pattern of feedback control as the learning task but involves change in direction or space displacement of some one component. Also, negative transfer may be found if the learning and transfer tasks are organized similarly overall, but require a shift from simple to more complex patterns of feedback control, from less precise to more precise control patterns, or from nondelayed to delayed control or vice versa. Many examples of negative transfer in machine skills appear to involve these factors.

An unsolved problem of particular relevance to psychomotor skill training is that of asymmetrical transfer between two given tasks, where the problem is to identify the factor or factors which determine optimal order of training. In some cases, it appears that transfer is greater from a difficult task to an easier task than vice versa, but this is not always true (Holding, 1962). For example, in tracking tasks where difficulty is assessed by accuracy scores, transfer usually is better from the more precise task to the less precise. Our own analyses of skill learning lead us to believe that asymmetrical transfer effects are related to the complexity of the feedback control patterns, where complexity may involve the spatial organization of an external task pattern, the intrinsic organization of movement components, or the type and degree of feedback transformations involved. It appears that transfer generally is better from more complex to less complex tasks than vice versa.

## Receptor Cybernetics and Perceptual Learning

In the present cybernetic view, all aspects of the receptor or input process —i.e., sensitivity, receptor orientation, pattern recognition, sensory quality, adaptation, and time functions—are regulated by interacting feedback-controlled neuromotor systems which generate specific movements of the receptor or sensory surface. Thus, perception at any stage of learning reflects the nature and refinement of these cybernetic control mechanisms.

Observations and research findings extending back to Cajal (1911) provide evidence that the receptor systems of the body are neurally organized for feedback control. Cajal observed that nerve fibers arising in visual centers of the brain extend back to the retina. Such efferent-afferent connections also have been identified in the auditory system (Davis, 1951) and in the proprioceptive system (Granit, 1955). Centrifugal efferent-afferent connections also have been found in the pyramidal motor system feeding back to the dorsal column nuclei of the cat's brain (Magni, Melzack, Moruzzi, & Smith, 1959). According to our cybernetic interpretation, these efferent-afferent internuncial cells govern the sensitivity and pattern of stimulation of the receptor surfaces and the afferent input into motor centers by comparing differences between the efferent output beyond the receptor and the afferent input to the receptor.

For decades it has been recognized that certain specific movement mechanisms govern the stimulus process of receptor systems. Prior to World War II, systematic analyses of optokinetic tracking in the guinea pig, cat, and man disclosed that these dynamic movements of the eyes are differentiated relative to level of evolution of focal vision (K. U. Smith & Bojar, 1938), relative to the development of vision and visual acuity in ontogeny (Warkentin & Smith, 1937), relative to the neural levels controlling pattern vision, visual acuity, and movement vision (K. U. Smith, 1937, 1939, 1940, 1941; K. U. Smith & Bridgman, 1943; K. U. Smith & Kappauf, 1940; K. U. Smith, Kappauf, & Bojar, 1940), relative to directional neural control mechanisms of the superior colliculus (K. U. Smith & Bridgman, 1943), and relative to a number of extinction, fatigue, inhibitory, and learning processes and their neural basis (K. U. Smith, 1941). Since the concept of cybernetic control of visual tracking was clarified during the war period, a major challenge has been to extend the experimental study of vision as a response-controlled afferent process.

More recent findings provide evidence that the retina's sensitivity is specifically dependent on its dynamic properties as a feedback-control system. Research with the stabilized retinal image (Fender, 1964; Pritchard, 1961; Riggs & Ratliff, 1951) has demonstrated that the eyes' tremor and flick movements are essential for maintaining retinal sensitivity and pattern vision. This type of research complements other findings long recognized as indicating that the accommodative and pupillary mechanisms of vision and the tympanic system of the middle ear are closed-loop mechanisms which regulate the visual and auditory afferent processes. Also, it is widely recognized that there are distinctive autonomic neurohormonal feedback circuits which affect the state of the pupil and the accommodative mechanisms as well as the vascular bed of the retina.

We regard research of this sort on stimulus stabilization mechanisms as just the beginning of a cybernetic analysis of the receptor process in learning and adaptation. We may ask what happens to the input functions of receptors when the afferent process of a particular receptor system is tied in a closed-loop arrangement with its own motor system and the feedback properties of this loop are then varied. The main objective of experimental receptor cybernetics is to develop methods for the study of variation in and perturbation of these afferent feedback-control systems. In our own laboratory, we have explored techniques of introducing spatial, temporal, and kinetic variations in eye-movement feedback, using optical, television, and closed-loop computer systems.

## CLOSED-LOOP COMPUTER CONTROL OF THE EYE

Special optical devices have been used to carry out cybernetic analyses of the eye as a control system and to study the learning functions with variable feedback control (Fender, 1964; Fender & Nye, 1961; W. M. Smith, 1964). Our own effort has been to perfect closed-loop computer techniques not only to analyze time sampling of feedback, as Young & Stark (1963) have done, but also for systematic experimental variation of all significant properties of the feedback-control process (Smith, *et al.*, 1963a). As indicated in the third diagram of Fig. 2, we arrange a closed-loop digital computer system with an eye-movement recorder and an oscilloscope target display so as to establish closed-loop control of optical movements with the oscilloscope target as a feedback image. The eye movements are recorded photoelectronically and the analog signal is amplified and converted to digital form. The digitized eye-movement signal then is delayed or otherwise perturbed by the computer, deconverted into analog form, and used to activate the oscilloscope display.

A critical part of this technique is to use computer programing to correct for nonlinearities in eye-movement recording relative to the magnitude of target movement. That is, the computer is used to make the oscilloscope spot move proportionately to the extent of the eye movements. After the subject is fitted with the eye-movement recorder and arranged in a fixed head position, he is instructed to watch the movement of the target blip. The computer then is used to generate five successive target spots and to measure the eye movements as the target jumps from one position to the next, establishing the relationship as an output target-control function. Thereafter, as the subject moves his eyes—for example, to one of the positions indicated—the target will move to the same position. Accuracy of such computer control has been refined to about .5 degree. Computer programing also is used to schedule trials and to recalibrate before each trial.

During experimental observations, subjects are instructed to try to keep the target blip centered on a marked spot on the oscilloscope. A typical procedure then is to introduce a spatial perturbation such as a sine wave into the feedback image and also to delay transmission of the image. In this situation, individual subjects show marked variation in how they attempt to establish control over the perturbed visual image. Some subjects produce systematic compensatory movements while others use both compensatory corrective actions and positive tracking movements in attempting to follow the delayed image and bring it to the center point. Still other subjects seem to make a spatial correction for the delay effect and in this way correct to some extent for the target displacement. In all cases, the error effects vary with the magnitude of perturbation of the feedback target and with the magnitude of delay. The effect of practice is to reduce to some extent the magnitude of the error in feedback control, but the degree of learning depends very greatly upon the particular modes of response control adopted by the subject. Only limited adaptation can be made to the delayed feedback condition inasmuch as the target position always will lag behind the eye movement. Although subjects learn readily to correct for spatial perturbations alone, no amount of practice enables them to make precise corrections for spatial perturbations while the feedback image is delayed. Subjects appear to adapt less successfully to delayed feedback from eye movements than to that from body movements such as manipulations or speech where tactual and proprioceptive control can be substituted to some extent when exteroceptive feedback is delayed.

## CONTROLLED DISPLACEMENT OF EYE-MOVEMENT FEEDBACK

A promising step in vision cybernetics is the development of lens-prism systems which permit variable control of the spatial displacement of feedback from eye movements. A great deal of work, of course, has been done on the effects of displaced vision on general body movements, but such research tells us little or nothing about feedback control of the eye movements themselves. To achieve a more complete visual displacement than is effected by inverting and reverting spectacles, the writer has had built[3] a contact lens-prism system which can be placed directly on the eye, as shown in Fig. 12.

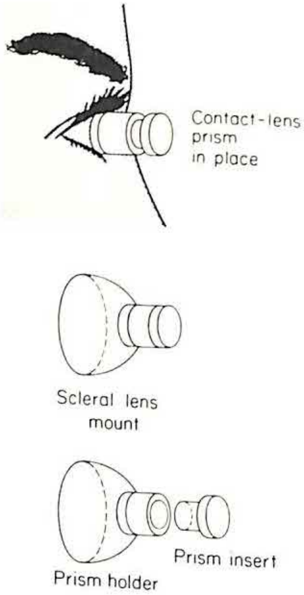


FIG. 12. Contact lens fitted with rotating prisms for study of geometric displacement of visual feedback from eye movements. (In collaboration with Charles Schick.)

---

[3] In collaboration with Charles Schick, Indiana University.

He wore this contact lens with its adjustable prism mount for about 1 hour before the observations were interrupted by technical difficulties. The visual effect was somewhat more disconcerting than wearing goggles fitted with inverting prisms, but it could not be determined whether this was due to the discomfort of the lens or to the more complete feedback displacement effected by mounting the prism directly on the eyeball. There was far greater instability of vision with the inverting contact lens than with inverting spectacles and there was no evidence of adaptive effects during the period that the lens was worn.

Successful control of the spatial properties of eye-movement feedback may be a major step in the analysis of the eye as a neurobehavioral control system. Hopefully, we shall be able to compare the effects of different kinds of displacements on feedback control of the optic system and to measure to what extent the visual system adapts to controlled feedback displacements. These results then can be compared with the effects of equivalent geometric displacements of feedback from head, hand, and gait motions, as produced by prism and television systems which do not directly affect eye-movement control.

Interest in cybernetic control of the perceptual process also has led to research on the specialized feedback control of reading. In one study mentioned earlier, the transfer potential of inverted reading training was tested in relation to writing with inverted feedback (Rhule & Smith, 1959). Although a highly uniform learning function was demonstrated for reading prism-inverted material, no transfer occurred when the task was changed to writing in the inverted field, thus showing that receptor cybernetic regulation is differentiated from visual-manual feedback control. In a series of studies on the feedback organization of reading behavior, it was demonstrated that reading responses display normal and breakdown ranges of angularly displaced feedback, and conform to principles of axial coordinate feedback regulation (K. U. Smith, Cambria, & Steffan, 1964b). Learning functions for spatially displaced reading material varied systematically relative to the degree and dimensions of displacement.

## PERCEPTUAL LEARNING AND RECEPTOR CYBERNETIC CONTROL

Although research on receptor cybernetics is just in its beginning stages, the field already constitutes a revolutionary reorientation to psychophysical analysis, perception, and learning. In the studies mentioned above, we dealt only with the problems of variable feedback control of the external ocular muscle systems of the eye. The equally significant problems of analyzing the internal ocular muscle system and the vascular network of the retina as cybernetic regulators of the retinal process have not yet been subjected to the cybernetic approach. In addition, the effects on vision and visual learning of physiological perturbations from the cardiac and respiratory systems, which no doubt regulate in part the afferent processes of different receptors (Sokolov, 1963), have as yet to be investigated.

Current innovations in research on receptor cybernetics represent a fitting centennial observance of Helmholtz's (1925) launching of the empirical doctrine of perception. However, current results promise to correct for some of Helmholtz's limitations of judgment in assuming that learning rather than genetically defined physiological factors plays the primary role in determining the characteristics of space vision. Cybernetic research indicates decisively that factors of intrinsic physiological specialization and division of labor in reflex action have a primary function in the visual process, and that learning—the empirical factor, so-called—has a derivative and dependent role. The nature of the feedback process as it is perturbed both by external and internal relationships and events determines the primary sensitizing, alerting, orienting, oscillatory, searching, and vigilance characteristics of sensory or perceptual response. Adaptation, extinction, and learned refinement of the movement-controlled feedback processes can occur with repeated stimulation, depending on the time characteristics of the repeated stimuli. One main effect of learning is to refine and stabilize the afferent process by reducing physiological perturbations. Also, the postural and transport components become less prominent with practice as the perceptual activity becomes more localized and more refined. Finally, the system compensates to some degree for persisting feedback perturbations, thus maximizing perceptual discriminations and resolution of stimulus patterns. The image-forming process of vision is thus an integrated cybernetic interaction which involves both intrinsic feedback loops and their modification by learning. This account of the refinement of space perception through learning does not conform with Helmholtz's interpretation of a century ago. Neither does it conform with current assumptions of the empirical origins of perception based on conditioning and reinforcement concepts.

## Cybernetic Design and Instrumental Learning

The cybernetic account of instrumental learning extends the principles of closed-loop feedback control to situations characterized by feedback transformations imposed by the instruments or machines being used. Any tool or instrument, such as an airplane, a lathe, or a bifocal lens, effects systematic transformations in the space, time, force, velocity, or intermittency characteristics of sensory feedback and thus modifies the pattern of response control and the nature of the learning function.

Although there has been some recognition among psychologists that a tool or machine or language or other symbol system transforms the control of response and its learning, there have been no systematic concepts prior to the development of cybernetic theory to specify the nature of such transformations. The cybernetic assumption is that machines and symbolic systems interpose systematic additions to and alternations of the closed-loop feedback-control processes between the motor or output phases of reaction and the sensory-feedback input. We recognize three main sources of stimulus feedback from instrumental behavior that the individual may utilize to govern his movements. These include feedback from movements of a bodily part, such as the arm or hand, feedback from the motions of the tool or machine, and feedback from the action of the tool or machine on environmental materials or objects. We refer to these three types of feedback respectively as reactive, instrumental, and operational feedback. In our view, the use of verbal language and nonverbal symbolism also involves sources of reactive, instrumental, and operational feedback.

We make these assumptions about reactive, instrumental, and operational feedback in tool and machine operations: (1) that these types of feedback are independently variable but interrelated in determining response and learning; (2) that various complementary, compensatory, and interfering interactions can occur between different sources of feedback; (3) that instrumental operations often involve variable concurrent and consecutive patterns of delayed feedback which interact with the spatial factors to determine the pattern of closed-loop control; (4) that the various sources of feedback produce both dynamic and static patterns which are used in closed-loop control; (5) that tool-using operations such as cutting, shearing, smashing, pushing, pressing, lifting, throwing, carrying, piercing, forming, and directing are defined specifically by sensory patterns produced by interaction of instrumental and operational sources of feedback and are in no way dependent on reinforcement after-effects; (6) that complexity and difficulty of a particular machine or tool operation are determined by number, variability, precision, and magnitude of the different sources of feedback transformation that characterize the design of that particular tool or machine; and (7) that the effectiveness of tools and machines for learning and efficient operations is determined by how well their design integrates the different kinds of feedback.

## Experimental Cybernetic Research on Instrumental Learning

To test some of the assumptions listed above, we have devised television techniques of differentiating reactive, instrumental, and operational visual feedback in order to measure the effects of each on learning. One study used a dual closed-circuit television setup to record both the front and the back of a "magic slate" panel. A front camera televised the movements of the hand and a marking or tracing tool, while a back camera televised the operational effect of the performance—i.e., the marking, writing, or drawing trace. The images from the two cameras could be presented separately or combined on a single monitor.

Another technique for separating different feedback sources is to use contrasting illumination effects in a television display. For example, if the hand or parts of a tool are blackened, these parts are invisible if televised against a black background. In one study, a set of black marking pencils was prepared with limited white areas so that each pencil provided visible feedback from a different locus—from the writing tip, from the cone, from the shank, or from a narrow band located some 3 centimeters from the tip.

The most general finding of our studies of differential feedback is that instrumental performance and learning are most efficient when feedback is provided from the locus of instrumental operation. In a writing or tracing task, the individual performs best if he can see the writing point or the written trace or—preferably—both. In writing, as in many other simple instrumental tasks, however, the movement of the hand conforms generally to the movement of the tool, so that reactive feedback, instrumental feedback, and operational feedback have a high degree of congruence. When reactive feedback differs in its spatial organization from instrumental or operational feedback, or both, the individual is more highly dependent on the latter sources. For example, a lathe operator achieves precision not by watching his hand but by watching the action of the cutting tool against the material being turned.

## DYNAMIC AND STATIC FEEDBACK

A problem significant to general learning theory concerns the relative effectiveness of dynamic and static feedback in instrumental behavior. By dynamic feedback we mean the stimuli resulting from actual bodily movements or the actions of a tool or machine. By static feedback we mean the nonmoving, persisting results or after-effects of movements or operations, such as the trace left by a pencil or typewriter in writing. Either type of feedback may function as knowledge of results, but we do not make the gratuitous assumption often made in learning research that feedback also functions as reinforcement. Feedback is an effect which is related systematically to the activity in progress, whereas reinforcement—insofar as it can be defined at all—is an effect extrinsic to the activity, which must be related to a drive or motivating state in order to be effective. According to cybernetic principles, tool-using tasks will be learned far more efficiently with dynamic feedback of performance than with only the static after-effects of response, even though the latter provides knowledge of results.

We used the delayed-feedback technique to make an experimental comparison of dynamic and static feedback, for static after-effects often by nature are delayed until a task or a segment of response has been completed. Closed-circuit television provided both kinds of feedback, either a static image of the operational effects of performance displayed on the television monitor at the conclusion of a task, or a dynamic image of the entire performance recorded on video tape and displayed after the task. It should be noted that neither of these conditions is optimal for effective performance and learning, for in each case the subject performs "blind" and must wait until after the task for knowledge feedback.

The task was to dot 12 small target spots with a felt marking pen in a prescribed manner. Subjects who received dynamic feedback saw a delayed image of their hand dotting with the marking pen, while those who received static feedback saw only the pattern of dots left at the end of their performance. In addition, some of the subjects were shown normally oriented television images, while others were shown inverted images.

Learning functions for the four experimental groups are shown in Fig. 13. Variance due both to learning and to feedback conditions was significant. Performance was more accurate with dynamic than with static feedback, and more accurate with upright than with inverted feedback.

These studies of tool-using performance are limited in number but represent a promising new line of attack on instrumental learning. The preliminary results indicate that the different kinds and sources of feedback produced in instrumental tasks have differential effects on the level of performance and the course of learning. These results support our general theory that specialized transformations of feedback-control patterns are involved in instrumental and machine behavior, and that the course of learning of such behavior varies with such factors as spatial complexity, temporal and kinetic variations, and other properties of the stimulus patterns functioning in closed-loop regulation of response.

The instrumental learning with which we are concerned here is in no way equivalent to the type of animal learning sometimes referred to as "instrumental," for a rat or monkey making a direct response to some surface feature or cue of a complex man-made situation is not using tools in the sense that men use tools. Further, human tool using and machine performance are not explained by learning theories that reduce the so-called instrumental learning of animals to open-loop associative, conditioning, or reinforcement learning models. Such theories do not account for the differential sources of sention that differs in many important respects from much psychological thinking, including conventional S-R theories of learning. Although cybernetic theory is based on the concept of feedback control, our interpretation of feedback is not equivalent to certain ideas that can be found in the learning literature. The feedback principle that grew out of wartime tracking research has found considerable acceptance among psychologists because they have assumed that feedback is no more nor less than the familiar "knowledge of results" factor of learning economy. However, it is further assumed by many learning theorists that knowledge of results functions as reward or reinforcement as well as information. For example, feedback has been described as a form of reinforcement contingency which the individual may anticipate as a perceptual effect (Adams, 1961). These interpretations do not emphasize the positive aspects of the feedback signal, but consider it as an error signal which presumably is reinforcing in inverse proportion to its magnitude. Although there have been some efforts to distinguish among different functions or roles of feedback or among different kinds of feedback (Bilodeau & Bilodeau, 1961), there has been very little effort to divorce the feedback concept from the ubiquitous reinforcement of contemporary learning theory.

The present cybernetic view is that although different forms and sources of feedback can be differentiated, including the more or less immediate dynamic sensory effects of movement as well as static after-effects which function as knowledge of results, feedback should in no case be equated with the poorly defined concept of reinforcement. We consider an analysis of the variable closed-loop factors of response as the starting point in an understanding of learning and believe that such a cybernetic approach is incompatible with conventional open-loop learning models. This conclusion is based on the many distinctions that can be drawn between cybernetic analyses and conventional association theories of learning. Conventional theories have no concepts with which to deal with the processes and relationships analyzed in behavioral cybernetics, and to attempt to preempt feedback as a form of reinforcement is to ignore the conceptual and experimentally defined differences between cybernetic and conventional association theories. The principal differences are indicated below.

### Specification of Learning Variables

Closed-loop factors which control learning can be defined both quantitatively and qualitatively, whereas reinforcement and associative factors cannot be defined in objective terms that have general validity. We can specify different modes of sensory feedback—visual, auditory, tactual, and kinesthetic—as well as measurable variations in spatial, temporal, and kinetic patterns.

Different conditions of dynamic and static feedback can be defined as well as different sources of reactive, instrumental, and operational feedback in tool using and machine operation. In contrast, definitions of reinforcement are tautological, for they are based on the very processes they are supposed to explain. No one knows whether a given extrinsic condition or event will be reinforcing in human learning until the fact has been demonstrated to provide hindsight.

## Spatial Organization of Response

Directionality and spatial organization are characteristic of all animal and human behavior but cannot be dealt with at all in the association and reinforcement models, which assume temporal factors as the basis of response organization. Directional guidance is especially noteworthy in human sensorimotor skills, which never have been handled satisfactorily by learning theory. Directional orientation of behavior has had to be assumed by learning theorists even though they cannot explain it. For example, Pavlov (1927) assumed that orientation behavior is a significant precursor to conditioning (Sokolov, 1963). Cybernetic theory provides a language and a method of analysis to describe the spatial factors in motor response, pattern perception, and neural control, in terms of the concepts of coordinate geometric feedback regulation and angular displacement of feedback in closed-loop guidance of response. Spatially controlled guidance mechanisms determine learning by serving as the bases of feedback integration of compound movements as behavior patterns develop and change in the human system.

## Temporal Organization of Response

Conventional learning theories have had to assume a flexibility in temporal organization of the significant stimulus and response events in learning. Different temporal arrangements are used in classic and operant conditioning, and learning still occurs even with wide variations from the optimal patterns. Such theories have no concepts to deal with the disturbances that occur as a result of time lags between movements and their sensory feedback. Cybernetic theory assumes that all properties of the feedback-control process are important in the determination of learning, and accepts research on delayed feedback as showing that response patterns are time specific. Temporal variations cause marked disturbance of the feedback control of response, prevent effective learning, and introduce marked inconsistencies in performance characteristic of feedback deprivation.

## INTERMITTENCY EFFECTS

One sector of learning research has dealt extensively with the effects of intermittent reinforcement on the strength of learned reactions, with results indicating that partial reinforcement maintains a higher level of response strength than consistent reinforcement. This finding in itself indicates the incompatibility between the reinforcement concept and sensory feedback, for there is no evidence whatsoever that intermittency in feedback contributes to stability of response. To the contrary, observations on the effects of intermittent feedback in tracking performance show that such intermittency reduces the efficiency of performance and learning (Searle, 1951). The cybernetic concept of time-sampled feedback has nothing in common with the learning theorist's concept of partial reinforcement.

## WITHDRAWAL EFFECTS

When the critical closed-loop sensory signals of a self-governing system are withdrawn, the response system ceases to operate and to learn. Severe and prolonged sensory deprivation in human and animal infants can impair development. In contrast, when reinforcements are withheld in animal conditioning experiments, we get experimental extinction—a gradual falling off in strength of learned response not comparable with the effects of feedback withdrawal.

The experimental extinction function has been given considerable importance in learning theory, although the phenomenon can be demonstrated in only a very few human situations. Many learned responses in man, such as verbal behavior and simple skilled movements, are repeated countless times with no identifiable reinforcement, yet no apparent extinction. When experimental extinction does occur in human learning experiments, it seems to be related to limitations in closed-loop self-regulation of response or to failure in pattern recognition and detection in the closed-loop system. That is, experimental extinction occurs most consistently in human subjects in restricted situations in which the individual exercises only the most limited control over the particular response under observation. Conditioned physiological reflexes may show the extinction function whereas somatic responses ordinarily do not. The uniformity of the extinction function appears to be related inversely to the degree of closed-loop regulation of the response system. Significantly, experimental extinction becomes less regular in subjects high in the phyletic scale, which possess more refined and diverse feedback-control mechanisms than do lower animals.

### Receptor Control

Since the first observations were made on action of the pupils, on eye movements in vision, and on responses of the tympanic muscles of the ear, it has been apparent that the input sensory processes of seeing and hearing are in part response controlled. Receptor cybernetic analyses have extended our knowledge of movement-regulated afferent input to include a variety of feedback mechanisms, e.g., movement control of receptor cell sensitivity and adaptation, somatic control of receptor orientation and pattern recognition, psychophysiological perturbation of the afferent process, general neurohormonal control of sensory mechanisms, and neural feedback patterning of the receptor signal itself. These cybernetic mechanisms regulate sensory input and thus help to define the S of learning theory's S-R model. Yet conventional theory has nothing to say about intrinsic cybernetic regulation of afferent events.

### Pattern Recognition

In cybernetic theory, pattern recognition is a necessary precursor of learning change. Prior to learning, the animal or individual must be able to detect direction and form in space and to control movements relative to such form recognition. Practice and experience serve to refine the primary modes of feedback detection and to extend normal angular displacement ranges and other feedback transformations that can be incorporated into control patterns. Man's ability to recognize and to remember countless different forms and patterns is derived from the unlimited number of spatial feedback relationships which can occur between motor and receptor surfaces.

The signal and spatial characteristics of stimulation on which both perception and perceptual learning depend are direct products of feedback-regulated motor response. This dynamic regulation of the receptor process is not just binary information handling, as specified by information theory, but involves multidimensional response patterns, including alerting reactions, postural control, receptor orientation, and articulated receptor action in discrimination. Space-organized coordinate feedback systems integrate the different components of receptor action to achieve precision of perception and to determine perceptual learning. In effect, receptor cybernetics constitutes a new science of the sensory process which describes in feedback language and analyzes with closed-loop techniques how the stimulus properties of the environment are processed by movement to produce recognized stimulus patterns.

The view that behavior development as well as learned adaptation depends on closed-loop signal processing stands in opposition to the empirical doctrine that space perception is altogether a function of experience. The cybernetic interpretation is supported by experiments on sensory-feedback deprivation and space displacement of vision as well as evidence concerning response control of the afferent process and the pattern of afferent excitation of the brain.

## GENERATION OF RESPONSE

Conventional treatments of learning do not explain the origins of the nonrepetitive responses which the learner brings to a new situation. Whether they be called trial and error, random activity, or emitted responses, such "spontaneous" reactions are recognized as a necessary prelude to learning, but learning theory offers no clue as to how they are generated and organized. In fact, the traditional view that all organized responses must have been learned through past experience seems to negate the possibility of reacting flexibly to new situations.

The feedback concept of learning proceeds from the basic postulate that a behaving system is to some degree self-regulating and self-generating. Its natural state is activity, not equilibrium. Thus, we say that the behaving system is not homeostatic but homeokinetic. The directed activity which is an essential feature of a closed-loop cybernetic system is generated by the system itself, and in this sense behavior is self-motivating. Energy is mobilized and expended independently of the physiological drive states which traditionally are assumed to underlie all behavior and learning.

The implication of these ideas is that learning is not dependent on general physiological drives, anxiety states, or unconscious frustrations as assumed by the various formulations of reinforcement theory. Assumptions about the drive-learning interaction are derived for the most part from animal studies (especially rat psychology), which have utilized hunger, thirst, pain, and other fairly predictable motivating conditions to activate their arbitrarily confined and restricted subjects. To explain behavior that seemingly is unrelated to physiological drives, the learning theorist—rather than abandon his drive-reduction model—has postulated that the individual is driven by exploratory, perceptual, or achievement motives, second-order or derived motives. The feedback concept offers learning science an alternative to the exigencies of the drive-reduction and reinforcement models by postulating that the behaving system is designed to generate its own responses and to regulate them continuously according to spatial relationships between the body and the stimulating environment.

### MULTIDIMENSIONALITY OF RESPONSE

One of the main findings of our cybernetic research is that response is organized in terms of independently variable coordinate feedback mechanisms related to posture, transport, and manipulative movements. Multidimensional response control also has been shown to be a central aspect of behavioral development in infants. Traditional learning psychology, based as it is on generic concepts of response, has no basis for conceptualizing the multidimensional nature of response and the differential roles played by movement components in development, learning, and instrumental and symbolic activities.

### SPECIALIZATION IN LEARNING

A number of systematic studies have yielded convincing evidence that learned responses are specialized in terms of the infinitely variable spatial properties of their feedback-control patterns. Related studies of delayed sensory feedback prove that such specialization cannot be based on flexible temporal patterning of stimulus and response as is indicated by traditional association theories.

The cybernetic view that learned specialization of response is the underlying factor which precedes and defines memory and transfer is in accordance with findings on transfer and learning interaction dating from the earliest studies. In contrast to association theories of learning, which have offered no verifiable basis for response specialization, feedback research has shown that specialization in learning, memory, and transfer can be based on the spatial properties of the controlling feedback process.

### FEEDBACK TRANSFORMATIONS

Cybernetic theory assumes that feedback regulation is based primarily on built-in relationships (spatial, temporal, and kinetic) between the motor and sensory systems and that variations in closed-loop response patterning will be produced by variations and transformations of the primary relationships.

Unaided movements of the body generate feedback stimuli with spatial, temporal, and kinetic properties which correlate closely with those of the generating movements. When tools and instruments are used, instrumental and operational feedback stimuli are generated in addition to those produced by movement itself. Thus, the tool or machine transforms cybernetic control by creating instrumental and operational feedback patterns that differ from those of unaided movement, but which still bear certain systematic spatial, temporal, and kinetic relationships to the form and dynamics of the controlling movements. Performance efficiency and learning thus depend on the design of the tool – i.e., upon its degree of correlation with the intrinsic feedback properties of the response system concerned.

The multidimensional feedback interactions which permit the countless variations of response patterning in instrumental behavior also make possible the infinitely variable symbolic transformations of control, either nonverbal or verbal. Systematic symbols such as drawings or words are used by individuals or groups in place of the primary more direct sensorimotor patterns to control particular adaptive responses. Cybernetic theory of learning assumes that symbolic responses are related to direct sensorimotor actions and instrumental behavior in terms of systematic transformations of the feedback properties of the more basic forms of response. However, because of the systematic evolution of language to highly complex forms, verbal behavior achieves a high degree of independence from the primary forms of psychomotor behavior in the learning of the individual.

The symbolic patterns involved in language and educational skills have specific designs related to their reactive, instrumental, and operational feedback properties. Their design determines how languages and symbols can serve as modes of self-regulation and self-stimulation in learning, as instrumental tools for communication, and as operations for transforming information in thinking and problem solving. The shift from direct to symbolic behavior may involve simple substitution of one response pattern for another or it may consist of complicated operations of reducing, translating, elaborating, and converting symbols which represent situational patterns of behavior.

This transformational concept of symbolic behavior assumes that symbolic learning and direct reactive learning follow the same general principles. Thus, the more complex or precise the transformational feedback involved in a particular symbolic or language process, the greater the difficulty in learning control of the process and its applications. The simpler the symbolic process, the easier it is to learn. It should be possible to specify the complexity of symbolic processes and languages, including mathematics and artificial languages, in terms of the number, kinds, complexity, and precision of the transformations of reactive, instrumental, and operational feedback involved in them. No matter how complex such transformations may be, however, patterns of symbolic control still bear systematic relationships to fundamental patterns of receptor and motor control from which they derive their meaning.

This account of instrumental and symbolic behavior points up the principal distinction between cybernetic theory and association or reinforcement theory of the determining factors in learning. In the cybernetic description of response and learning, the nearly unlimited capacities of the human response mechanisms to generate systematic transformations of feedback control provide the objective basis for human specialization. Human response patterns and learning capacities are assumed to have become more refined and more complex during the course of man's biosocial evolution. In contrast, conventional association and reinforcement theories propose that each individual in turn must acquire through fortuitous contiguities or reinforcement contingencies all of the complicated instrumental and symbolic skills characteristic of his culture. In our opinion, human skills are too consistent, too systematic, and too complicated in their patterns of control to be attributed to the hazards of temporal contiguities, the chances of reinforcement contingencies, or the vicissitudes of undifferentiated drive states in the life of each individual.

### Types of Learning

Learning investigators have found it necessary to distinguish among a great variety of learning types, such as classical conditioning, operant conditioning, verbal rote learning, psychomotor learning, problem solving, concept formation, and others. A significant fact about these categories of learning is that they cannot be derived from the association or reinforcement learning models which are accepted by the same investigators who propose the classifications. Rather, the so-called types of learning refer to differences in the intrinsic stimulus-response factors involved—to the way stimuli are controlled by continuous self-regulated reactions. The need for creating such classificatory schemes reflects a fundamental failure in current learning theories which cannot be rectified by labeling the various procedures used in learning experiments.

From a cybernetic point of view, the different forms or variations in learning reflect differences in patterns of feedback control which the animal or human subject can utilize in a particular learning situation. Variations can occur in the modes of response required for closed-loop regulation, in the sensory channels involved, or in the spatial, temporal, and kinetic properties of the feedback processes governing response. Learning functions also can vary because of spatial, instrumental, or symbolic transformations which alter the complexity and form of feedback control. In this view, different types of learning result from variations in the modes, conditions, and transformations of feedback stimuli from postural, transport, manipulative, and receptor movements involved in different learning situations.

The distinction we are making here between conventional and cybernetic interpretations of learning types is of greatest significance in relation to problems of human education and training. The fact is that conventional association theories incorporate no generally valid concepts of educational design. The so-called categories of learning which have originated in laboratory designs for learning research have generated theoretical arguments for decades but have provided no systematic understanding of the sensorimotor factors which define behavioral control in human educational or training situations. In contrast, cybernetic theory postulates that all of the various learning categories represent different conditions and transformations of feedback control, and that human training and education can be designed according to the same general cybernetic principles that govern all instances of learning change. Cybernetic theory proposes that the high degree of precision and efficiency achieved in speech, in writing, in reading, in mathematics, in technology, and in artistic production be attributed not to the chance occurrence of contiguous events or vaguely defined rewards and punishments, but to the objectively defined and infinitely variable sensory-feedback patterns in terms of which the individual controls his own responses and environmental events.

## REFERENCES

Adams, J. A. Human tracking behavior. *Psychol. Bull.*, 1961, **58**, 55–79.

Bilodeau, E. A., & Bilodeau, I. McD. Motor-skills learning. *Annu. Rev. Psychol.*, 1961, **12**, 243–280.

Black, J. W. The effect of delayed side-tone upon vocal rate and intensity. *J. speech hear. Disord.*, 1951, **16**, 56–60.

Black, J. W. The persistence of the effects of delayed side-tone. *J. speech hear. Disord.*, 1955, **20**, 65–68.

Brown, G. G. Perception of depth with disoriented vision. *Brit. J. Psychol.*, 1928, **19**, 117–146.

Cajal, S. Ramon y. *Histologie du système nerveux de l'homme et des vertébrés.* Vol. 2 (Transl. by L. Azoulay). Paris: Maloine, 1911.

Chase, R. A. *An information flow model of the organization of motor activity.* Washington, D.C.: Neurocommunications Unit, National Institutes of Mental Health, 1963.

Conklin, J. E. Effect of control lag on performance in a tracking task. *J. exp. Psychol.*, 1957, **53**, 261–268.

Davis, H. Psychophysiology of hearing and deafness. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 1116–1142.

Duncan, D. B. Multiple range and multiple F tests. *Biometrics*, 1955, **11**, 1–42.

Ewert, P. H. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior. *Genet. Psychol. Monogr.*, 1930, **7**, 177–363.

Fairbanks, G. Systematic research in experimental phonetics: I. A theory of the speech mechanism as a servosystem. *J. speech hear. Disord.*, 1954, **19**, 133–139.

Fairbanks, G. Selective vocal effects of delayed auditory feedback. *J. speech hear. Disord.*, 1955, **20**, 333–346.

Fairbanks, G., & Guttman, N. Effects of delayed auditory feedback upon articulation. *J. speech hear. Res.*, 1958, **1**, 12–22.

Fender, D. H. Control mechanisms of the eye. *Sci. Amer.*, 1964, **211** (1), 22–35.

Fender, D. H., & Nye, P. An investigation of the mechanisms of eye movement control. *Kybernetik*, 1961, **1**, 81–88.

Fitts, P. M. Engineering psychology and equipment design. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1287–1340.

Gebhard, J. W. *Some experiments with the VF aided tracking equipment*. Appl. Phys. Lab., Johns Hopkins Univer., Rep. No. 166-1-53, 1948.

Goldiamond, I., Atkinson, C. J., & Bilger, R. C. Stabilization of behavior and prolonged exposure to delayed auditory feedback. *Science*, 1962, **135**, 437–438.

Gould, J., & Smith, K. U. Angular displacement of visual feedback in motion and learning. *Percept. mot. Skills*, 1963, **17**, 699–710.

Granit, R. *Receptors and sensory perception*. New Haven, Conn.: Yale Univer. Press, 1955.

Held, R., & Hein, A. V. Adaptation of disarranged hand-eye coordination contingent upon re-afferent stimulation. *Percept. mot. Skills*, 1958, **8**, 87–90.

Held, R., & Schlank, M. Adaptation to disarranged eye-hand coordination in the distance-dimension. *Amer. J. Psychol.*, 1959, **72**, 603–605.

Helmholtz, H. L. V. von. *Physiological optics*. Vol. 3 (Transl. from 3rd German ed., J. P. C. Southall [Ed.]). Rochester, N.Y.: Opt. Soc. Amer., 1925.

Holding, D. H. Transfer between difficult and easy tasks. *Brit. J. Psychol.*, 1962, **53**, 397–407.

Hubel, D. H. The visual cortex of the brain. *Sci. Amer.*, 1963, **209** (5), 55–62.

Hubel, D. H., & Wiesel, T. N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *J. Physiol.*, Lond., 1962, **160**, 106–154.

Hull, C. L. *Principles of behavior*. New York: Appleton, 1943.

James, W. *Principles of psychology*. Vol. 1. New York: Holt, 1890.

Kohler, I. Über Aufbau und Wandlungen der Wahrnehmungswelt, insbesondere über 'bedingte Empfingungen.' *Sber. Öst. Akad. Wiss., Phil.-hist. Klasse*, 1951, **227** (1) (Transl. available by G. Krauthamer). (a)

Kohler, I. Warum sehen wir aufrecht? *Die Pyramide*, 1951, **2**, 30–33. (b)

Kohler, I. Experiments with prolonged optical distortion. *Acta Psychol.*, 1955, **11**, 176–178.

Lee, B. S. Effects of delayed speech feedback. *J. acoust. Soc. Amer.*, 1950, **22**, 824–826. (a)

Lee, B. S. Some effects of side-tone delay. *J. acoust. Soc. Amer.*, 1950, **22**, 639–640. (b)

Lee, B. S. Artificial stutter. *J. speech hear. Disord.*, 1951, **16**, 53–55.

Levine, M. Tracking performance as a function of exponential delay between the control and display. *USAF WADC tech. Rep.*, 1953, No. 53-236.

Licklider, J. C. R., & Miller, G. A. The perception of speech. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 1040–1074.

Lincoln, R. S., & Smith, K. U. Transfer of training in tracking performance at different target speeds. *J. appl. Psychol.*, 1951, **35**, 358–362.

McDermid, C., & Smith, K. U. Compensatory reaction to angularly displaced visual feedback in behavior. *J. appl. Psychol.*, 1964, **48**, 63–68.

Magni, F., Melzack, R., Moruzzi, G., & Smith, C. J. Direct pyramidal influences on the dorsal-column nuclei. *Arch. Ital. Biol.*, 1959, **97**, 357–377.

Müller, G. E., & Pilzecker. S. Experimentelle Beiträge zur Lehre vom Gedächtniss. Z. *Psychol.*, 1900, Suppl. 1, 1–288.

Pavlov, I. P. *Conditioned reflexes.* (Transl. and ed. by G. V. Anrep.) London: Oxford Univer. Press, 1927.

Peterson, J., & Peterson, J. K. Does practice with inverting lenses make vision normal? *Psychol. Monogr.*, 1938, **50**, No. 5 (Whole No. 225). Pp. 12–37.

Phillips, R. S. *Aided tracking.* Radiation Lab., Rep. No. 453. Cambridge, Mass.: M.I.T., 1943.

Pritchard, R. M. Stabilized images on the retina. *Sci. Amer.*, 1961, **204**(6), 72–78.

Rhule, W., & Smith, K. U. Effect of visual pretraining in inverted reading on perceptual-motor performance in inverted visual fields. *Percept. mot. Skills*, 1959, **9**, 327–331.

Riggs, L. A., & Ratliff, F. Visual acuity and the normal tremor of the eyes. *Science*, 1951, **114**, 17–18.

Searle, L. V. Psychological studies of tracking behavior: IV, the intermittency hypothesis as a basis for predicting optimum aided-tracking time constants. *USN Res. Lab. Rep.*, 1951, No. 3872.

Siipola, E. M. Studies in mirror drawing. *Psychol. Monogr.*, 1935, **46**, No. 6 (Whole No. 210). Pp. 66–77.

Skinner, B. F. The generic nature of the concepts of stimulus and response. *J. gen. Psychol.*, 1935, **12**, 40–65.

Skinner, B. F. *The behavior of organisms: an experimental analysis.* New York: Appleton, 1938.

Skinner, B. F. *Verbal behavior.* New York: Appleton, 1957.

Smith, K. U. The relation between visual acuity and the optic projection centers of the brain. *Science*, 1937, **86**, 564–565.

Smith, K. U. The behavior of decorticate guinea pigs. *J. comp. Psychol.*, 1939, **27**, 433–447.

Smith, K. U. The neural centers concerned in the mediation of apparent movement vision. *J. exp. Psychol.*, 1940, **26**, 443–466.

Smith, K. U. Experiments on the neural basis of movement vision. *J. exp. Psychol.*, 1941, **28**, 199–216.

Smith, K. U. The geometry of human motion and its neural foundations: I, perceptual and motor adaptation to displaced vision. *Amer. J. phys. Med.*, 1961, **40**, 71–87. (a)

Smith, K. U. The geometry of human motion and its neural foundations: II, neurogeometric theory and its experimental basis. *Amer. J. phys. Med.*, 1961, **40**, 109–129. (b)

Smith, K. U. *Delayed sensory feedback and behavior.* Philadelphia, Pa.: Saunders, 1962.

Smith, K. U. Sensory feedback analysis in medical research: I, delayed sensory feedback in behavior and neural function. *Amer. J. phys. Med.*, 1963, **42**, 228–262. (a)

Smith, K. U. Sensory feedback analysis of visual behavior: a new theoretical-experimental foundation of physiological optics. *Amer. J. Optom.*, 1963, **40**, 365–417. (b)

Smith, K. U. Sensory feedback analysis in medical research. II. Spatial organization of neurobehavioral systems. *Amer. J. phys. Med.*, 1964, **43**, 49–84.

Smith, K. U., & Bojar, S. The nature of optokinetic reactions in mammals and their significance in the experimental analysis of the neural mechanisms of visual functions. *Psychol. Bull.*, 1938, **35**, 193–219.

Smith, K. U., & Bridgman, M. The neural mechanisms of movement vision and optic nystagmus. *J. exp. Psychol.*, 1943, 33, 165–187.

Smith, K. U., & Kappauf, W. E. A neurological study of apparent movement. *J. gen. Psychol.*, 1940, 23, 315–327.

Smith, K. U., & Smith, M. F. *Cybernetic principles of learning and educational design.* New York: Holt, 1966.

Smith, K. U., & Smith, W. M. *Perception and motion: an analysis of space-structured behavior.* Philadelphia, Pa.: Saunders, 1962.

Smith, K. U., Kappauf, W. E., & Bojar, S. The functions of the visual cortex in optic nystagmus at different velocities of movement in the visual field. *J. gen. Psychol., 1940,* **22**, 341–357.

Smith, K. U., Mysziewski, M., Mergen, M., & Koehler, J. Computer systems control of delayed auditory feedback. *Percept. mot. Skills*, 1963, **17**, 343–354. (a)

Smith, K. U., Wargo, L., Jones, R., & Smith, W. M. Delayed and space-displaced sensory feedback and learning. *Percept. mot. Skills*, 1963, **16**, 781–796. (b)

Smith, K. U., Ansell, S. D., Koehler, J., & Servos, G. H. Digital computer system for dynamic analysis of speech and sound feedback mechanisms. *J. Ass. comput. Mach.*, 1964, **11**, 240–251. (a)

Smith, K. U., Cambria, R., & Steffan, J. Sensory-feedback analysis of reading. *J. appl. Psychol.*, 1964, **48**, 275–286. (b)

Smith, W. M. Control of eye fixation by auditory feedback. *Psychon. Sci.*, 1964, **1**, 233–234.

Smith, W. M., Smith, K. U., Stanley, R., & Harley, W. Analysis of performance in televised visual fields: preliminary report. *Percept. mot. Skills*, 1956, **6**, 195–198.

Snyder, F. W., & Pronko, N. H. *Vision with spatial inversion.* Wichita, Kans.: Univer. of Wichita Press, 1952.

Sokolov, Y. N. *Perception and the conditioned reflex.* London: Macmillan, 1963.

Stratton, G. M. Some preliminary experiments in vision without inversion of the retinal image. *Psychol. Rev.*, 1896, **3**, 611–617.

Stratton, G. M. Vision without inversion of the retinal image. *Psychol. Rev.*, 1897, **4**, 341–360 and 463–481.

Stratton, G. M. The spatial harmony of touch and sight. *Mind*, 1899, **8**, 492–505.

Thorndike, E. L. *Educational psychology.* Vol. 2. *The psychology of learning.* New York: Teachers College, Columbia Univer., 1913.

Tiffany, W. R., & Hanley, C. N. Adaptation to delayed side-tone. *J. speech hear. Disord.*, 1956, **21**, 164–172.

Warkentin, J., & Smith, K. U. The development of visual acuity in the cat. *J. genet. Psychol.*, 1937, **50**, 371–399.

Warrick, M. J. Effect of transmission-type control lags on tracking accuracy. *USAF AMC tech. Rep.*, 1949, No. 5916.

Watson, J. B. *Psychology from the standpoint of a behaviorist.* (2nd ed.). Philadelphia, Pa.: Lippincott, 1924.

Wiener, N. *Cybernetics.* New York: Wiley, 1948.

Winchester, R. A., Gibbons, E. W., & Krebs, D. F. Adaptation to sustained delayed side-tone. *J. speech hear. Disord.*, 1959, **24**, 25–28.

Wolfe, J. B. The effect of delayed reward upon learning in the white rat. *J. comp. Psychol.*, 1934, **17**, 1–21.

Wooster, M. Certain factors in the development of a new spatial co-ordination. *Psychol. Monogr.*, 1923, **32**, No. 4 (Whole No. 146).

Yates, A. J. Recent empirical and theoretical approaches to the experimental manipulation of speech in normal subjects and in stammerers. *Behav. Res. Ther.*, 1963, **1**, 95–119.

Young, L. R., & Stark, L. Variable feedback experiments testing a sampled data model for eye tracking movements. *IEEE Trans.*, 1963, **HFE-4** (1), 38–51.

# Cybernetic Theory and Analysis of Learning

## Comments on Professor Smith's Paper

FRED ATTNEAVE
*University of Oregon*

In the preceding chapter, Smith has summarized a really imposing body of research and thinking on feedback and adaptation to distorted feedback. His paper has covered a broad range, and he has made a great many points, but three points impress me as particularly salient, interesting, and important, and I begin by considering these.

The first is the finding that, for each of several varieties of spatially distorted feedback, there is a "normal range" of displacement over which adaptation occurs rather readily, and a "breakdown range" over which adaptation occurs poorly or not at all. The most dramatic illustration of this distinction is in Smith's Fig. 8 (p. 449), which shows errors in maze tracing after long practice as a function of angular displacement of feedback. The error curve is practically flat up to 40°, and then rises sharply; i.e., performance breaks down abruptly with greater displacements. I doubt that many psychologists would have predicted this finding on the basis of either earlier results or common sense, though it may be pointed out (as Smith suggested in his use of the term *normal range*) that the normal angular motility of the arm, relative to the visual axes, perhaps has a range of about ±40°. The phenomenon leads us to believe that we are dealing with an analog control system into which a corrective constant can be introduced, but that whatever variable is thus corrected can be biased over only a limited range.

Second, I am impressed with Smith's hypothesis that there are three different control systems, differing in adaptability and possibly other respects, having to do with (a) vertical, postural, gravity-oriented activity; (b) horizontal, "transport" activity; and (c) manipulatory activity. The most striking empirical findings here are that inversion of feedback impairs performance on tasks like maze following much more than does left-right reversal, and that this differential impairment applies to the speed of left-right movements as well as to the speed of up-down movements. I still do not quite understand the reason for the latter result, granting the difference between horizontal and vertical control systems. Also, I have a couple of other questions about these experiments. Apparently both the motions and the feedback display occurred in a vertical plane. Suppose they were horizontally oriented instead, as on a table top. What would the prediction be in this case? Are movements away from and toward the subject, which correspond to up-down changes in the visual field, within the province of the vertical system or the horizontal system? If (as seems rather likely) the "same" results were obtained in a horizontal plane as in a vertical plane, one might conclude that the control systems are input-oriented rather than output-oriented, though other interpretations are no doubt possible. It is, moreover, somewhat unclear to me exactly what kinds of task are supposed to be controlled by the manipulatory system, and in what sense or to what degree such tasks could be independent of the first two systems. Finally, I wonder whether different control systems might be inferred from factor-analytic studies of motor skills, of the sort that Fleishman has performed.

Smith's third point that seems salient to me is that in continuous tasks people do not adapt in any real sense to temporally displaced feedback; that delayed feedback may at best be ignored, but not utilized. He feels that these results provide a strong basis for rejecting both contiguity and S-R reinforcement learning theories. I am no more an S-R psychologist than he is; nevertheless I find some of his arguments on this basis extremely elusive.

In the first place, it seems to me that a good part of the performance decrement resulting from delayed feedback may be explained in common-sense terms that are very nearly atheoretical. When feedback is "immediate," small errors can be corrected quickly, whereas with delayed feedback small errors can become large before the subject knows they are occurring. This is a performance effect, which will operate at any level of learning, whatever the nature of the learning. I agree completely with Smith that what is most often called a *response*—e.g., an act like sitting down—is in fact a highly complex sequence of motions, continuously guided by feedback, and that the slipperiness of the *response* concept (closely paralleled by the slipperiness of the *stimulus* concept) is no credit to the S-R position. The S-R theorist may respond to this criticism in either of two ways. He may say that an act is made up of many micro responses, segments of behavior that could, at least for analytical purposes, be considered shorter than the feedback time and accordingly responsive to feedback. The lack of any obvious natural boundaries for such micro responses should be clearly noted. On the other hand, he may say that sitting down and similar acts constitute functional units of behavior, that they are reinforcible in a unitary manner, etc., and that he will therefore deal with them as responses, feedback and all. This position is an uncomfortable one for several reasons: the "response" is rather a plan or program the detailed execution of which is conditional upon feedback; moreover, explaining how it *attains* a status of functional unity is itself a legitimate task of learning theory.

Over and beyond these considerations, however, Smith is saying, I believe, that any learning theory based on temporal contiguity would predict more learning than actually occurs under conditions of delayed feedback. It is this argument that I have trouble following, and that I hope he will spell out.

In arguing more specifically that feedback is not to be identified with *reinforcement* (or *knowledge of results*), he pointed out that Wolfe's rats showed considerable learning in a T maze when reinforcement was delayed for 30 seconds, or even for several minutes, whereas much shorter feedback delays in human maze tracing and the like do not produce learning. However fundamentally misguided the reinforcement theorist may be, in this case he can argue back, quite legitimately, I believe, that the possibilities for *interference* are vastly different from one of these situations to the other. If the rat's reward is delayed for 10 seconds, it does not matter if his behavior in the delay box is reinforced even more than his behavior at the choice point. But if a human subject is tracing a star maze with feedback delayed for 10 seconds, a great deal of interference or maladaptive learning will occur if the "reinforcement" at each instant operates more strongly on behavior more recent than 10 seconds past. I am not contending, and I do not believe, that feedback is "merely" reinforcement. Feedback does not merely tell the subject that he is right or wrong; it also tells him *how* he is wrong. I do believe, however, that the difficulty of learning (as opposed to performance) under conditions of feedback delay may be attributed to some kind of associative interference.[1] Irrespective of psychological theory, certain information-processing operations appear to be absolutely necessary prerequisites to the occurrence of learning under delay conditions: (a) temporary storage of traces of the subject's behavior; (b) appropriate (rather than inappropriate) matching or referencing of these traces to corresponding segments of feedback when the latter arrive. Does any theory actually assume or predict that people can carry out these operations under conditions of continuous performance?

---

[1] After writing these comments, I read the paper of Dr. Ina McD. Bilodeau in this volume and found that she has presented detailed, thoroughly convincing evidence for the view that interference is the major factor here. She also discusses the separate effects of feedback delay on learning and on performance.

Although I do not agree with everything Smith has said, I agree with most of it—insofar as I understand it—and my subsequent commentary takes the exceedingly annoying course of inquiring why he has not dealt with things that I would have discussed if I had been writing his paper. Both parts of his title, "Cybernetic Theory and Analysis of Learning," established in me some expectations that were never quite fulfilled.

In the first place, I think of cybernetic theory as dealing in some detail with the formal properties of information and control systems. It is concerned with what the variables of the system are, what functions relate these variables to one another, and with the consequences of these functional relationships. Where feedback is involved, the exact nature of the feedback is important. (To cite a well-known example, a simple single-loop negative feedback system may do one of three qualitatively different things, depending on its time constants and other parameters: it may proceed to a steady state either monotonically or via damped oscillation, or it may oscillate indefinitely.)

It appears, at least superficially, that Smith is unconcerned with this level of analysis. For the most part he deals with time and error scores, which discard the detailed information about the subject's performance that would presumably be necessary for inferences about the formal properties of the control system. Consider for a moment a moth that is set, at a given time, to maintain a particular angular orientation to the sun as it flies, and suppose that the angle of orientation has a value between 0° and 90°. Now substitute a candle for the sun, and the moth will fly a roughly spiral course into the flame, as moths do. If a human subject attempts to move a stylus from one point in an open field to another, with angular displacement of visual feedback and with no compensation for the displacement, the simplest prediction is that he will follow a spiral course, exactly like the moth. I doubt very much that a real subject behaves in any such simple way, but why doesn't he? Would it be legitimate, for example, to interpret deviations from the predicted spiral as evidence for immediate adaptation?

It may well be that the performances of individual subjects in tasks of the kind under consideration are so complex as to defy analysis on this level at the present time, and that Smith considers such concerns premature. My examples are undoubtedly too simple; as multiple feedback loops are added, the conceptual complexity of a system increases very rapidly indeed. (A known case of multiple-loop negative feedback at quite a low level is the limulus eye, in which each element inhibits, and is inhibited by, its neighbors. It is a pretty safe bet that motor control systems involve a much greater complexity than this.) I strongly suspect that such complex systems are beyond the practical range of ordinary analytical procedures, and that computer simulation—in which the program embodies the theory—offers the only real hope for dealing with them. Smith's reaction to these comments will be of great interest to me. What I should like to ask him more generally, I suppose, is how he envisions the long-range future development of behavioral cybernetics. In what terms, and with what degree of precision, will we be able to explain and understand the systems with which he is dealing in another 50 years?

"Analysis of Learning" is the other half of Smith's title. I am quite in agreement with his negative appraisal of S-R theory, though for reasons that may not be exactly the same as his. What puzzles and frustrates me, however, is that he has not really stated what kind of alternative theory he espouses. I have read a good many, though not all, of his other works, and I am still puzzled.

We find such tantalizing hints at a theory as: "...the major conclusion that should be drawn from the research on delayed feedback is that disturbances induced in behavior and learning are the result of perturbation of the immediate movement-produced sensory-feedback stimuli. These stimuli are absolutely essential for closed-loop response control and learning..." (p. 446). This generalization would have seemed very rash indeed a few years ago, but the research of Held and his associates, as well as that of Smith, offers some impressive support for it. Still it does not constitute a theory, as I understand that term, because it leaves the formal nature of the learning process unexplained. Elsewhere (p. 429) Smith said, "Thus, the reflex is conceptualized not as an S-R unit but as a pattern of feedback-mediated interactions between the motor and sensory systems which provide a measure of behavioral control." I wonder if this is intended to suggest that the reflex is acceptable as a unit of learning, provided it is redefined to include feedback components.

To quote further: "The cybernetic theory proposed here suggests that the brain be conceptualized not as a synaptic system but as a neuronic system. Internuncial cells function as neuronal units in detecting differences in movement-produced feedback stimuli" (p. 429). Again: "Cybernetic theory implies directional guidance of response based on detection of differences in stimulus patterns" (p. 447). Here I have several questions to raise. First, are the postulated detectors sensitive merely to differences in the present input, or may they also detect differences between the present input and some template, or schema, or neuronal model? It seems to me that goal-directed behavior may require something like the latter function. Second, when only the present input is involved, just what are the elements between which the differences are detected? In the case of perceptual analyzers that we actually know something about, e.g., the slant detectors of Hubel and Wiesel, we can say at the very least that they are not merely responsive to differences between pairs of cones. Third, and most relevant to our present concern, how does the proposed system learn? If learning does not occur at the synapse, where and how does it occur?

I find it interesting and peculiar that Smith has referred often to learning theorists with whom he disagrees, like Watson, Thorndike, Skinner, and Hull, but has given no recognition whatever to theorists whose ideas might be more compatible with his own, namely, Tolman, Hebb, Miller, Galanter and Pribram, and Chomsky. I wonder if he has not taken the essentially Tolmanian position that the organism learns paths in a cognitive space, i.e., behavioral rules for transforming one input into another. It should be remarked that Hebb's phase sequence clearly takes feedback into account, and that feedback is an absolutely essential aspect of the "TOTE unit" of Miller, Galanter, and Pribrám, whose orientation is frankly Tolmanian. Certainly Chomsky's transformational grammar, and his concern with the ability of people to utter original sentences, is highly congruent with Smith's thinking.

Now, I may well be overinterpreting. It is perfectly true that I have only a glimmering, at best, of what Smith's learning theory is, and I hope he will explain it. It would clarify my thinking a great deal if he would take some very simple, but representative case of learning, and draw a block diagram, or flow chart, or similar representation showing the variables that result in learning, the locus or loci of the structural changes they effect, and the principles by which they effect these changes.

# Chapter 10

# Motor-Skills Learning and Verbal Learning: Some Observations

BENTON J. UNDERWOOD
*Northwestern University*

The observations recorded in this chapter are a varied, perhaps even ragged, lot. At one extreme they are concerned with general systematic issues, and at the other with a detailed examination of certain measurement problems. If there is an underlying theme it is that of examining certain aspects of the current approaches to research in motor-skills learning and in verbal learning, but, as will be seen, there are exceptions to this theme. Certain trends in the contemporary research in the two fields are noted but not emphasized. To assert the existence of trends is (inferentially, at least) to project into the future. The trouble with this is that the future may show one to be quite wrong. What looks like a trend today may turn out to be a fad or a bust, and the whole matter is given over to the oblivion which we tactfully call our archives.

Each reader may assess the future of motor-skills research for himself by examining what the experts in the field are doing and thinking as set down in the previous chapters. These investigators are the ones who will give the thrust to the research in the area and thus shape the future. The trends we note in verbal learning will probably also tell us something about the future. But let us hope they do not tell us too much. How dull it would be if we could not approach each new experiment with an expectation that we might discover something quite new and thus help mold the future: indeed how dull, and how quickly most of us would change fields if we believed it true.

To place motor-skills learning and verbal learning in a conjunctive relationship may seem to exact a promise, a promise of a unifying theme or of a search for common concepts in the two areas. However, this was not the central purpose of the conference and it is not the central purpose of this chapter. The search for common concepts and common phonomena across the areas constituting the field of learning was the purpose of a previous conference (Melton, 1964a). The fact is that some of the participants in the present conference have done significant work in both skills learning and verbal learning and it was inevitable that certain parallels be drawn. A parallel method of exposition will be used in the present chapter but largely for the purpose of organization.

The problems of delimiting an area of learning that might be called motor-skills learning are thoroughly discussed by Irion (Chapter 1). When motor-skills learning is spoken of here we mean the learning observed on any of a great variety of tasks, from those requiring very simple positioning movements to those requiring complex and coordinated tracking skill. Many of these tasks are described in the previous chapters and no purpose is to be served by repeating these descriptions here. In the same manner we may identify the area of verbal learning by specifying the learning behavior observed when certain tasks are presented the subject. Thus, we think of the paired-associate task, the serial-learning task, verbal-discrimination tasks, and so on. Selective-learning tasks (Chapter 2) have certain properties that do not fit easily into either the motor-skills category or the verbal-learning category but may, at the same time, have some common dimensions with both. In the discussion which follows, the emphasis is on the more distinctly motor-skills tasks such as represented by tracking, and the verbal-learning area is normally typified by the paired-associate task.

The superb history of motor-skills research provided by Irion has no counterpart in verbal learning to which we may refer. As a poor substitute for such a history of verbal learning, some comments are made initially about the work in verbal learning in the past which may serve as a crude basis for evaluating the changes which have taken place in recent years.

As we all know, historically in the United States verbal-learning research was closely associated with that loosely knit set of attitudes known as functionalism. Experimental reports from Columbia University or the University of Chicago, or from students trained at these two schools, constituted the bulk of the persistent work on verbal learning. To simplify and to exaggerate somewhat, this work may be said to be characterized by two features. First, gross independent variables were manipulated to determine functional relationships between these variables and learning and retention. From such studies, empirical generalizations were developed if the data from experiment to experiment showed enough consistency to warrant the generalizations. Second, there was but little original detailed theoretical content in the work on verbal learning. Except in a most general way, the independent variables chosen for manipulation were not chosen because of their theoretical relevance. A variable such as similarity was manipulated because it was known to influence interference, but the detailed way in which interference operated was seldom considered.

For a few years prior to World War II and for a few years after, two quite different research cultures, both dealing with learning, existed side by side in some laboratories. In the animal laboratory, variables were being specified by extremely detailed theoretical considerations as the Tolman, the Hull, and the Guthrie positions were arrayed against each other. In the verbal-learning laboratories, with certain exceptions, the theoretical component was minimal. There were no good theoretical guys and bad theoretical guys; the sustenance of research in verbal learning was provided by the discovery of empirical regularities, not by decisions about theories.

The various disturbances produced by World War II did not allow an assessment of the direction the research in verbal learning was taking until about 1950. At that time it became quite apparent that it wasn't taking any direction; it was relatively stagnant because so few people were working in the area. Shortly thereafter, however, the area began to share the increase in research which characterized most of psychology as well as other disciplines. It is not this increase as such which is of moment here, but rather the features of the research which began to take form and which today characterize the methods of attack on the problems in verbal learning. Although the development of the features which characterize research in verbal learning today seems almost inevitable (if other areas are used as models), there had to be some beginnings. To the present writer, two events occurring in 1940 established the precedent for the change in the approach to verbal learning which eventually took place, and which to a large extent dissociated verbal learning from the classical functionalist attitudes.

One of these events was Gibson's (1940) theoretical analysis of certain phenomena in verbal learning. The critical aspect of this theory (in terms of its influence on later research workers) was the manner in which it conceived of the gross learning of a verbal list as being constituted by a set of subphenomena, with each of these subphenomena having some empirical status in its own right. As was customary at that time at Yale University, most of the subphenomena came from the conditioning laboratory. No attempt of this type on so broad a scale had ever been undertaken previously in verbal learning, and the approach was to have considerable impact.

The second event which was to influence subsequent research workers in verbal learning was the postulation of unlearning as a factor in retroactive inhibition (Melton & Irwin, 1940). This notion, at the time it was introduced, was strictly a hypothetical construct (initially it was called Factor X). Since it was introduced by one who still calls himself a functionalist (Melton, 1963), it can only be concluded that the introduction of such a concept in 1940 was quite out of keeping with the behavior of the classical functionalist. In any event, the ties between verbal learning and the functionalism of the past were considerably loosened by this unfunctionalist act.

With these brief historical remarks concerning verbal learning, we may turn to some observations of the current scene in verbal learning and in motor-skills learning.

## Unit of Analysis

### VERBAL LEARNING

The basic unit of analysis in verbal learning is an association. The notion of an association was bequeathed to us from pre-experimental thought and also became the unit of analysis for verbal-learning studies in the functionalists' laboratories. However, contemporary conceptions of an association may hold little resemblance to the older conceptions, and because those not familiar with current research in verbal learning often confuse the two, it will be well to provide some characterization of current conceptions.

(1) The paired associate is often spoken of as consisting of a stimulus (S) term and a response (R) term. It thereby is commonly identified with S-R psychology as a theoretical position. This may or may not be the appropriate identification, depending upon the investigator. No one is so naïve as to believe that an association is formed between two verbal units apart from the responses the subject makes to these units. It is, therefore, quite possible to become an R-R theorist, although many prefer to think of the response to the stimulus term as having stimulus properties which become associated with the response produced by the formal response term.

(2) We know that if the stimulus term consists of more than one element, the response which the subject comes to make to this term (hence, the response which enters into the association) may be less than the compound presented to the subject. This is spoken of as stimulus selection, and one of the active areas of research in verbal learning at the moment is that of determining the variables which govern stimulus selection (e.g., Cohen & Musgrave, 1964). It is simply not reflecting current thought to visualize an association as being etched in a one-to-one correspondence between what is presented the subject and what he learns.

(3) The association which develops, and which we call learning, may be a multistep or chaining affair. When the subject learns to produce B when A is shown, it does not imply necessarily that there is a simple direct association between the representations of A and B. The work on mediation (e.g., Jenkins, 1963) makes it evident that intermediaries may occur. The subject brings to any learning situation a vast network of already formed associations and these may enter into the formation of a so-called new association.

These associative networks may be mapped (e.g., Deese, 1964), and they enter into theoretical formulations of both learning and retention.

(4) The acquisition of an association is influenced by the context in which it appears. Lists of pairs are commonly used and we know that learning of a given associate is influenced by the relationships among the associates. The experimental study of intralist similarity is a major way of examining the influence of context, although it is not the only variable which defines contextual factors.

(5) An association acquired in paired-associate learning has bidirectional properties. The drawing of an arrow from S to R is a mechanical matter to represent the task given the subject in the usual paired-associate experiment. The R term may become a stimulus to determine the presence of the association from R to S (backward association). It has been argued by some (Asch & Ebenholtz, 1962) that the two different directional properties of an association cannot be distinguished on the basis of strength.

(6) An association is a functional unit and simply reflects the observation that under certain specified conditions one event will be followed by another—that learning will occur. The notion of an association is tied neither to a switchboard conception of the nervous system nor to any other neurophysiological conception.

The above six points represent some of the major features of modern associationism as it is being molded in the verbal-learning laboratory around paired-associate learning. All tasks used in the study of verbal learning, however, do not yield as well to the S-R approach as does paired-associate learning. In free learning, where units are presented in different orders on successive trials with the subject recalling in any order he wishes, the stimulus event for the recall of any given unit is difficult to specify. A puzzling matter has arisen with regard to serial learning, although the fact that it is puzzling may result only from a long-held conception that serial learning results in a series of S-R associations. While it may consist of such a series of units, the facts are that contemporary research efforts have been generally unsuccessful in specifying what each successive stimulus event is (e.g., Young, 1962). It must be clear, therefore, that for certain verbal-learning tasks the notion of an association as a basic unit of analysis represents a certain faith that sooner or later appropriate experimental analysis will provide data which conform to the conception. It should be noted that the issue bridges both the empirical and the theoretical. On the empirical side, we may continue to do research with free learning and serial learning using as our response measure the number of units produced correctly *as if* discrete associations were involved. But if theoretical conceptions require specification of the precise stimulus events, a clear limitation is placed on the theorizing about such verbal-learning tasks.

## Motor-Skills Learning

As will be clear to anyone who has read the preceding chapters, the use of an association between a particular stimulus event and a particular response event as the unit of analysis for motor-skills learning has at best only very limited applicability. The approach is obviously quite useful in selective learning, and in fact, Birch's discussion of Chapter 2 includes a suggestion for inserting verbal mediating responses between the stimulus event and the response, much as is being done in verbal learning. Simple positioning movements might also be fitted into the associationistic framework. But when we view tasks which require continuous performance, such as tracking, the idea of discrete associations is difficult to maintain. As Irion has pointed out in his chapter on the history of motor skills, there is no substitute for or parallel concept to that of the discrete association which seems serviceable for the continuous task. If the elementary association is lacking as a unit of analysis, certain other characteristics of associations likewise do not have meaning for continuous tasks. For example, it is difficult to think of backward associations in a tracking task.

In verbal learning we sometimes speak of response integration. Thus, a three-letter trigram such as *PFK* is said to be difficult to integrate. In a certain sense, when it *is* integrated, it shows (in behavior) the same continuity as the behavior in a well-learned tracking performance. We often assume that the trigram is acquired by forming successive associations (P to F and F to K). But this trigram is in fact a three-item serial list, and we noted above the difficulties experienced thus far in attempting to specify particular associations in a serial list. Taken individually, each letter (P, F, and K) can be produced (by speaking) without difficulty—the motor skill required for these productions has been mastered. But the fact is that these three independent skills have rarely been practiced in this sequence, and other sequences are antagonistic to them. Putting them together is difficult. It is problematical just how much of this difficulty is to be attributed to a skill factor, but the evidence shows that it takes longer to spell a sequence of three letters which rarely occurs in our language than to spell a sequence of equal length which occurs frequently (Waters, 1939). We could, perhaps, imagine a tracking task consisting of three movements, each of which when considered individually was practiced to a high level of precision, but when we put them together the resulting sequence may be of the kind that has rarely been practiced. This analogy cannot be pushed very far, but it is possible that we will find eventually that acquiring a serial verbal list does have much in common with a continuous tracking task. In any event, neither type of task has thus far appeared amenable to analysis by elementary units of association.

There is another related point which has been made by Poulton (Chapter 8) which bears repeating. He pointed out that we rarely study the learning of the skill to execute a particular verbal response. Learning to give the proper pronunciation to words in a foreign language would be an illustration. Or suppose we required our subjects to respond with a particular frequency and intensity in paired-associate learning in addition to responding to the appropriate stimulus. Under these circumstances we would be more nearly studying skill learning. As it is, we merely use the vocal response as an instrumental response to demonstrate associative learning, and the vocal characteristics of the response that we will accept as correct are very broadly defined.

Of course, research has proceeded in spite of the inability to specify particular associations in many motor-skills tasks. Indeed, some would say it is a blessing not to be so bound. Stimulus variables can be manipulated and performance measures can be related to the manipulations. While the isolation of particular associations (S-R units) is not a theory as such, it almost necessarily dictates the formal characteristics of the theory which will be used. We may expect, therefore, that the theories that arise from data from skills tasks will be different from those that are prompted when the unit of analysis is the association. It is clear in this volume that this expectation is supported. Nevertheless, it is also apparent that there is something attractive to investigators when they *can* specify particular stimuli which are involved in producing particular responses.

## Constituent Analysis

### Verbal Learning

Much current work in verbal learning may be described as consisting heavily of constituent or component analysis. The legacy of the classical functionalists' work was a series of gross relationships. Much of the work in verbal learning of the past few years represents experimental analysis of these gross phenomena so that subphenomena or constituent processes are isolated. These constituent phenomena become, then, objects of study in their own right. Battig (Chapter 5) has shown how constituent analysis (component analysis) has proceeded in the case of key transfer paradigms. If one wishes to make a prediction of the transfer to be expected for a given paradigm there are at least a half-dozen constituent phenomena which must be considered in making the prediction. Each of these constituent phenomena may be studied independently by appropriate experimental manipulation.

A few further rather simple illustrations of constituent analysis may be mentioned. If one wishes to study relatively pure response learning (the acquisition of verbal units as units), free learning may be used. If one wishes to study the formation of associations between stimulus and response terms without the encumbrances of response learning, the subject may be given the response terms (say, the numbers 1 through 9) prior to learning. If learning from trial to trial is conceived of as a resultant of both acquisition and forgetting processes, an investigator may study the short-term retention of a simple paired associate with varying numbers of other associates interpolated, thus determining the amount of forgetting to be expected from one trial to another in paired-associate list learning. If one wishes to study the forgetting of various components of the learning of a paired-associate list, different forms of recall procedures may be used to isolate the amount of the forgetting of each component.

There is nothing new in constituent analysis. It typifies the method of approach which sooner or later is seen in almost any experimental discipline. It may be directed by theoretical analysis or by logical analysis. It involves a refinement in the control of independent variables and also a refinement in response measures, both of which indicate a simplification of the experimental task. If a single feature can be used to typify the stage of research in an area at any moment in time, we would typify verbal learning today as being in the stage of constituent analysis. Constituent analysis is relatively immune to widely different systematic points of view; it is being used extensively by those who bring a gestalt heritage to the laboratory, by those who reach their analyses via mathematical models, and by those who might be classed simply as modern behaviorists. Constituent analysis appears to be so characteristic an approach at a certain stage in the development of an area that it must tell us something about a basic commonality in the thought processes of investigators.

## MOTOR-SKILLS LEARNING

The thinking and research of workers in motor-skills learning, as typified in the previous chapters, also shows evidences of the use of constituent analysis, but there are some differences between this work and that discussed above for verbal learning. Among workers in verbal learning, there appears to be some unity of purpose in constituent analysis. Although constituent phenomena may be experimentally abstracted from a more complex task, such as paired-associate learning, and although these constituents may be studied themselves for what they are, there appears to be some consensus that fundamentally we are studying the constituents to help us understand (in the long run) the more complex task. This goal may be somewhat less apparent in the work on motor skills, but it is clearly present. We may illustrate this initially with some nonlearning constituents. There has been much recent work on the problem of reaction time, an aged area of research which until about 1950 had seemed moribund. But it became apparent that to understand the behavior exhibited in certain tasks it was important to know more about reaction time than was known. Some of this work, particularly the work relevant to tracking behavior, has been reviewed by Poulton (Chapter 8). With certain displays it has also become necessary to study aspects of perception, including search processes. And Adams (Chapter 4) has proposed to deal with attention in a systematic fashion in order to understand better the stimuli on a display to which the subject is in fact responding.

There is also considerable evidence of constituent-analysis thinking in the work on motor skills with the constituents being learning phenomena. If a track, such as a series of steps of different magnitudes, is presented the subject, his behavior clearly suggests a form of probability learning if the steps of different magnitude occur with different frequencies. This learning may be a distinctly different constituent in the overall performance than the learning of the actual skill of tracking. Each could be studied separately. Any experiment fitting the classical designation of part-whole learning is making an assumption about the constituent learnings involved in the overall learning.

In spite of the fact that we can find these evidences of constituent analysis in the motor-skills area, it is clearly not as pervasive an approach as it appears to be at the moment in verbal learning. We point out in the next section that some investigators seem to have reservations about the approach when viewed in light of what they wish to accomplish by their research. Constituent analysis may have two outcomes which, while not necessarily incompatible, may be so viewed. First, it is a means of discovering new phenomena, phenomena which may be described as being less complex than the gross phenomenon from which they were extracted. Second, constituent analysis may be viewed as a means of understanding the complex. It is with regard to this second outcome that differences in opinion are manifest, and we will consider these differences in the next section.

Three final points need to be made. First, constituent analysis seems to be more easily handled and leads more readily to a mutual understanding when elementary units of association can be used. Thus, the S-R unit may provide a common reference point for different tasks and for different investigators in the field of verbal learning. Insofar as skills learning does not have a commonly accepted unit of analysis, it is difficult to know when constituent analysis in the area has reached a common level of understanding for various tasks.

Second, it is worth noting that Battig (Chapter 5) makes a strong plea for the use of transfer designs in motor learning as a means of constituent analysis. This approach seeks not to study transfer as such, but to use the performance on the transfer task as an indicant of the nature of the learning which took place in the initial task. Finally, we should note that constituent analysis must sooner or later lead to the problem of synthesis, and it is to this problem that attention is now turned.

## Synthesis

### VERBAL LEARNING

Implicit in the discussion of constituent analysis in verbal learning was the assumption that having "torn" the learning of a paired-associate list apart in terms of the constituent processes, we could in some way sum these processes and get the whole. It is apparent that precise tests of the assumption will not be made until quantitative values or weights are assigned each constituent process and the combination rules worked out. Some advances toward this end are being made by those using mathematical models (e.g., Restle, 1964). But even without the virtues of precise quantification, some progress may be possible. This may be illustrated in a somewhat simplified version of a more extended discussion of the problems (Underwood, Ekstrand, & Keppel, 1965).

Assume we are studying the role of conceptual similarity in paired-associate learning. More particularly, we are interested in the role of conceptual similarity among stimulus terms vs. a corresponding similarity among response terms. All terms may be animal names for a high-similarity condition, and the words for the low-similarity condition would have minimal conceptual relationships. Of course, the words in the two kinds of lists would be equivalent on all other characteristics. Assume we have 10 pairs in each list, with the numbers from 1 to 10 being the neutral terms (e.g., when the words are stimuli the numbers will be responses, and vice versa). What predictions can be made?

From free-learning studies we know that a series of items which are conceptually related will be learned more quickly than an equal number which are not conceptually related. This fact is asserted to be relevant to the acquisition of responses in a paired-associate list. Thus, it must be predicted that high similarity will "throw in" a positive factor in over-all paired-associate learning when these high-similarity words are response terms. There is no counterpart to this positive effect when the words are the stimulus terms. We may assume for the moment that associative interference from high similarity will be equivalent when the words are response terms and when they are stimulus terms. This is to say, if we measured in pure form the difference in associative interference resulting from low and high similarity the slope and height of the curve would be exactly the same for response-term similarity as for stimulus-term similarity. Thus, in this simplified situation, the only factor which leads to differential predictions of the effect of stimulus similarity and response similarity is response learning. The predictions therefore become clear: (1) learning will be more rapid when the words are stimulus terms than when they are response terms. This follows from the fact that the numbers from 1 to 10 will be acquired more quickly (as response terms) than will words, regardless of the similarity among the words. This prediction is of no particular interest here. (2) The slope of the curve relating stimulus-term similarity and trials to learn the paired-associate list will be steeper than the one relating response-term similarity and trials to learn. The positive effect on response learning of high similarity among the response terms will "subtract" from the negative effect of associative interference produced by these response terms. Thus, the over-all critical prediction is an interaction between similarity and the locus of the words (stimuli or responses). The prediction has been confirmed in the study noted earlier.

This is a relatively simply illustration of the summation of constituent processes, but it is sufficient for present purposes. But the question may be raised as to the precision with which the summation can be made. As noted earlier, the only completely satisfactory way to handle this is through quantification of the constituents. However, when the situation is as simple as the one outlined, another approach may give us an appropriate answer and also give evidence on another problem which is always present in attempts to synthesize.

Assume we have determined the precise law between variations in similarity and the rate of response learning. That is, we have used graduated levels of similarity in different conditions of free learning, and have this law at hand. We now do the paired-associate study using these various levels of similarity. We then take the difference between the curve obtained when similarity was manipulated among the stimulus terms and the curve obtained from manipulating the same levels of similarity among the response terms. The slope of this resulting curve must be identical with the slope of the curve we obtained in relating free learning and similarity. It must be identical for we have assumed that the only variable factor as a function of the location of the words in the paired-associate list is difference in response learning as a function of similarity. If indeed the slopes are equivalent it would seem that our synthesis is handled appropriately. By looking only at the law for one constituent process we were able to appropriately predict the learning of the more complex task—the paired-associate task.

Assume we were not so fortunate; assume that we did not find equivalent slopes of the empirical and the derived curves. What are the alternatives? One is that our assumption about equivalent associative interference being produced by similarity among stimulus terms and among response terms is in error. As a test of this, we would try to devise a procedure which eliminated all factors except associative learning and see if the assumption is correct. If it isn't, this becomes a differential factor in our attempt to synthesize. Another alternative is that some unknown constituent process is involved. A third possibility is that our assumption of simple additivity is incorrect; the response-learning process within paired-associate learning may not follow exactly the same law as that found for free learning. Thus, to get a better estimate of the contribution of this component we may have to study response learning during the act of paired-associate learning. Finally, it will always remain a possibility that a given task, particularly one of some degree of complexity, has emergent properties. We may know all of the constituents and we may "add" these in many, many different ways and still not approximate the behavior on the complex task as it occurs on this task. Most of us would probably resist accepting this final alternative until it seemed absolutely necessary. For to accept it means that constituent analysis becomes a way of studying or discovering lower-level phenomena but these must become ends in themselves if we cannot show with any precision how they enter into the complex phenomenon from which they were extracted. This is not to minimize the importance of discovering these lower-level phenomena, and to identify the subphenomena making up a complex one does achieve some level of understanding of the complex. But in the long run we would be most pleased if the parts of the whole could be successfuly combined to produce the whole.

As pointed out earlier, the basic unit of analysis for paired-associate learning is the S-R unit. In adding constituent processes to obtain overall paired-associate learning we ask how each process will influence the learning of the S-R unit. It is probably correct to say that in terms of any index of absolute complexity the paired-associate list is not complex. We are capable of much more complicated forms of learning involving verbal material. If we consider synthesis in its broadest form, we may ask what our knowledge about paired-associate learning via S-R units will tell us about the ultimate in verbal learning, for example, linguistic behavior as exemplified in our abilities to generate (and to understand) new sentences. We have recently been given a very stern lecture on this matter (Miller, 1965). Miller's conclusion is that any attempt to understand language learning on the basis of S-R units and the variables influencing the acquisition of such units is doomed to failure. Although Miller's point of reference appears to be S-R analysis and theory as developed in the animal laboratory, his admonitions may be equally applicable to the verbal-learning laboratory. This is to say, what we know now and will learn in the future about simple associative learning may have sharply limited usefulness in attempts to understand the learning leading to the ability to assemble a group of verbal units into a new but meaningful sentence. However, some may remember Hull's use of habit-family hierarchies to handle the more complex behavior while still retaining the S-R unit. In the language of the computer theorist today, a similar notion is expressed by the use of subroutines and executive routines. Whether or not such modes of thinking can eventually handle the problems Miller raises must be left to the inscrutable future.

### MOTOR-SKILLS LEARNING

None of the participants in this conference was directly concerned with the matter of how constituent skills may be assembled or integrated to "add up to" the performance on a complex skill. Nevertheless, in the discussions attending the papers, and tangentially in certain of the prepared papers, the issue was touched upon. Obviously, before the problem of synthesis can be attacked empirically, analysis must have occurred. However, by the very nature of certain motor-skills tasks this analysis can be performed at a logical level without having to isolate experimentally the constituent behavioral processes. For example, if a task requires control of a rudder with the foot and a stick with the hand, we may ask about the acquisition of each skill by itself, and we may ask about the acquisition of the total skill involving both controls and whether or not it can be accounted for by a knowledge of the acquisition of the parts. Such an approach obviously suggests the manipulation of the classical variable, whole vs. part learning. Some of the studies on this variable in tracking tasks have been reviewed by Poulton (Chapter 8). For our purposes, the discussion will be directed toward an examination of the possible outcomes of experiments which manipulate the whole-part variable as these outcomes relate to the question of the assembly of the constituent skills.

Assume that a task is clearly divisible into two components, A and B. Each component can be practiced independently and skill can be developed on each by this independent practice. When the total task (A and B) is used, each of the independent skills will enter into the over-all performance, but the question is whether or not an additional behavioral component is added, perhaps, we might say, a coordination component.

One possible conclusion from a whole-part study, where the results show that the whole procedure is superior to the part procedure, is that the total task *is* more than the sum of A and B. Such a conclusion might be permissible if we can assure ourselves that in learning the parts only learning took place which was relevant to performance on the total task. This is related to the caution that has long been advised in such studies, namely, that in putting the parts together interference may occur. Thus, the superiority of the whole method of practice may result not from something extra in the whole that is not practiced in the part procedure, but because performance on the whole is reduced following part practice because of interference. This possibility may be illustrated by a brief return to verbal learning.

Assume the validity of the distinction between response learning and associative learning in a paired-associate task. To implement this distinction in a whole-part procedure, we might give one group of subjects initial practice in acquiring the response terms as such. Following this, the usual practice on the paired-associate task would be given. The performance would be compared with another group having had practice on the paired-associate task from the beginning. Such an experiment would probably show that the whole method was superior, but it could probably also be shown that this was due to interference resulting from part learning which was not present in the whole learning. In acquiring the response terms in part training, free learning or some variant on it would probably be used. Under such training procedures it is essentially impossible to eliminate the development of associations among the response terms. Such associations would subsequently interfere with paired-associate learning. This is to say, that while the subject had indeed acquired the response terms, he had also learned something that was incompatible with the requirements of paired-associate learning. In short, before whole-part studies showing superiority of the whole method can be adequately interpreted, we must be sure that in part learning only learning relevant to the total task occurred and that this occurred only to a level required for the total task. Only then do the results bear on the problem of synthesis.

If the outcome of a whole-part study shows no difference as a consequence of the training procedures, the conclusion would seem to be that we can "add up" the parts and get the whole. If the part procedure should prove superior to the whole, the interpretation would be quite obscure.

Another approach to the problem of synthesis is represented by attempts to predict performance on a complex task from the performance scores on the parts of the task. No attempt is made here to detail how these predictions are made. It is sufficient to say that the predictions have been only moderately accurate for relatively unskilled subjects, but quite accurate for skilled subjects (Bilodeau, 1957a, b). Insofar as such attempts are successful by simple assumptions of additivity, we are led toward the conclusion that synthesis may not be so difficult a matter in practice as it appears in the abstract.

It is not unusual to draw parallels between the flow of language and the flow of events in a well-practiced skill. Furthermore, we may generate a sentence on the spur of the moment which in its detail we have never produced before. In a similar fashion, perhaps, a professional basketball player may shoot a basket with grace and ease while holding a bodily position that he has never held before. We have seen that some believe that our understanding of simple associative learning will be of little value in attempting to understand language learning. By way of completing the analogy, therefore, it might be expected that some will maintain the position that the understanding of complex skills will be aided little by an understanding of simple skills. The approach used by Smith (Chapter 9) seems to accept this position. Furthermore, Smith has asserted that his theory is capable of handling not only the complexity of skilled performance but also the complexities of language learning. If this is true (and the present writer is quite incapable of judging the matter), Smith has presented a theory of extraordinary importance. Briggs, in his discussion of tracking, has presented a model designed to handle the learning of a complex tracking skill and it is apparent that he prefers to approach an understanding of this skill by working directly with it rather than by constituent analysis. There were suggestions in the discussion periods that others might accept a similar position.

It is clearly realized that the problem of synthesis accentuates a long-standing systematic issue. Some might prefer to call it a philosophical issue. It relates to the old question of whether the whole is greater than or different from the sum of its parts; it relates to the problem of whether phenomena may have emergent properties and cannot be understood in the usual sense by understanding the parts. The issue has been raised here for several reasons. First, it is patent that it is more than a philosophical issue for the position taken on it affects the strategies of research. Second, constituent analysis can only proceed so far at the behavioral level; sooner or later we reach a point where the behavior segment we are examining is at "rock bottom," at least with respect to available methods. It seems inescapable that more and more investigators are going to turn to the problem of putting the constituents back together in an effort to use the simple to explain the complex. Third, it seems that we can reach general agreement that whether or not synthesis is possible is after all an empirical problem. Much of our thinking, however, is geared to analysis; we may have to develop a new set of techniques to evaluate appropriately whether or not synthesis is possible in given situations.

## Theory

### VERBAL LEARNING

It was noted in a preceding section that verbal learning in earlier decades was closely associated with functionalism, and that functionalism was not known for its original theoretical contributions. The current work in verbal learning is associated with a variety of theoretical approaches, although no one person can boast of a comprehensive theory of verbal learning. Mathematical models have found a compatible companion in paired-associate learning. The computer has been used as a theoretical vehicle. Information theory or information analysis has been found useful by some in subsuming certain facts of verbal learning. Some of the work in immediate memory has been instigated by physiological models, and the notions of neural decay and neural consolidation are sometimes used as explanatory mechanisms. If diversity of theoretical approaches taps the heartiness of an area, the immediate future of verbal learning is bound to be a robust one.

It is probably correct to say that the typical approach to explanation of verbal-learning phenomena is not represented by any of the above-named modes of theorizing. The typical research worker prefers, it seems, to use explanatory concepts which have some independent empirical status. Many of these are still holdovers from the conditioning laboratory, concepts such as generalization, extinction, and spontaneous recovery. Some of these phenomena have been demonstrated to occur in the verbal-learning laboratory while others are assumed to occur, but in either case a part of the theorizing may be to assign the particular characteristics (how they are assumed to be influenced by particular independent variables) needed to account for certain findings. But there are also empirical concepts that are used which were discovered in the verbal-learning laboratory. For example, consider backward associations. These are phenomena for study in their own right, but they may also be used to explain certain previously puzzling phenomena such as the negative transfer which may occur in certain paradigms. Consider another illustration. If one assumes that word-association norms reflect the relative frequencies of *implicit* responses which are made to corresponding stimulus words in a verbal-learning experiment, then certain predictions about the learning in this experiment may follow.

The classical concepts of contiguity and frequency also occasionally appear in theorylike statements in verbal learning. Indeed, differences in assumptions about the role of frequency in the formation of associations are responsible for the lively contemporary controversy known as all-or-none vs. incremental learning. But, to repeat, if there is a modal manner of theorizing in verbal learning today, it is one which makes fairly heavy use of empirical concepts as explanatory concepts.

## MOTOR-SKILLS LEARNING

No attempt is made to review the various theories that have been applied to skills learning. No participant in this conference was given the nearly impossible task of reviewing theoretical approaches. Theoretical matters which did arise did so as an auxiliary outcome of the pursuit of other objectives. But even within these restrictions a great variety of theoretical approaches was discernible. This variety is even quite obvious in the theories dealing only with tracking behavior (Chapter 8). The present plan is to make brief comments about two theoretical notions, followed by a more extended treatment of a third.

No participant in the conference seemed impelled to defend the proposition that reinforcement in the classical sense should enter as a major theoretical notion into theories of skill learning. The analogy between the animal-learning laboratory and the skills laboratory can sometimes becomes quite close, as, for example, in the use of informative feedback (Chapter 6) in simple positioning movements. Nevertheless, the consensus seems to be that such feedback is best conceived of as a means of defining to the subject the nature of the response to be learned rather than as a reinforcing event in the classical sense. It is possible that the same conception might be applied in the animal laboratory.

The manipulation of temporal relationships among the events on a learning trial has accounted for much research in all areas of learning. It is difficult not to accept some principle of contiguity but theorists have differed as to which events must adhere to the contiguity principle in order for learning to occur. The heavy emphasis on temporal manipulations in the area of motor skills is evident throughout this book. The major events of concern are the response and the feedback to that response. The criticalness of this interval is not apparent in the types of informative feedback dealt with in Chapter 6, but it is apparent in other chapters where the feedback is concerned with the subject's perceiving the results of his responses as he attempts to control "something." In these cases the length of the delay between the response and the feedback becomes a highly potent variable.

In driving an automobile there is a relatively short lag between a steering movement and the "response" of the vehicle. In steering large ships, on the other hand, there is a long lag between steering movements and the response of the ship. In the laboratory, such differences may be simulated by the manipulation of lag time in tracking tasks. As the target or track changes the subject sees that he must react in a given way and moves a control to bring an indicator or cursor on the target. But, with a lag in the control system, he is not able to see the effect of his movement immediately. And of course, to stop the movement in a given direction once it has started he must make a correction appreciably before the display would seem to require it. Lags of even a few seconds can make the task extremely difficult to master. Of course, we have had much experience in controlling movements of simple tools and vehicles where the lag is essentially zero, and we may expect a disturbing effect when lag is present. It is also true that a lag requires the subject to learn precise temporal judgments, both as to the interval before any movement of his indicator occurs and the rate of change once movement is initiated, so that the lagged task may be intrinsically a more difficult one than a nonlagged task. Nevertheless, data from variation in lag times may be used in support of the thesis that contiguity between movements and the feedback from these movements is a critical variable in learning.

The role of contiguity in skill learning has perhaps been most dramatically shown in some of Smith's (Chapter 9) demonstrations. When even a slight delay is imposed between the subject's performing certain complex responses and the feedback, little if any learning may be observed. It becomes clear, therefore, that while Smith may rail against certain forms of contiguity theory, his data in fact must make him a staunch supporter of the contiguity principle.

Finally, we may turn to a consideration of a particular theoretical notion that has been closely associated with work in skills learning. The hypothetical construct of reactive inhibition (or temporary work inhibition) dominated one line of research in skills learning for a dozen years after World War II. The basic idea of a work inhibition (which builds up with responding and dissipates with rest) is an old one in psychology. But, as Irion pointed out in his history, the notion never gained prominence until Hull introduced it into his system. Subsequently it was taken over bodily by workers in the area of motor-skills learning. Indeed, in the initial phases of research on motor learning in the third period of history (as defined by Irion), the studies dealing with reactive inhibition must have contributed disproportionately to the increase in the research in motor skills in general, an increase also described by Irion. Equally dramatic is the description of how studies concerned with reactive inhibition have rather suddenly dropped to a trickle in the past few years. This lack of interest is reflected in this volume, where one finds only a passing reference to it in one paper (Chapter 8). Had this conference been held 10 years ago it would not have been surprising to have seen the conference revolving around inhibitory constructs, with reactive inhibition as the central point of reference. This loss of interest in reactive inhibition appears to the present writer to be an outstanding puzzle in the contemporary sociology of our discipline. For this reason, and for others which will subsequently become apparent, some examination of the various aspects of this puzzle will be undertaken.

Perhaps we need first to remind ourselves that reactive inhibition is a hypothetical construct, and if we speak of it sometimes as if it were an empirical phenomenon it will only demonstrate how commonly accepted the notion is. As a second reminder, we should note that the phenomena said to be primarily caused by reactive inhibition are of large magnitude. That performance measures for a task may actually decrease under massed practice while increasing under distributed practice attests to the power of the mechanism assumed to be responsible. When inferred from reminiscence scores following a single rest period subsequent to massed practice the shift in the performance level upward may be so dramatic that we may even allow our students to "get away" without using a control (no rest). These comments about the magnitude of the effects associated with reactive inhibition are made for two reasons. First, there are not many phenomena that we deal with in the laboratory (including learning itself) which are of such magnitude as those associated with reactive inhibition. Second, given this magnitude, we have an ideal situation for experimentation in that there is much "room" to play with. If there are constituent processes associated with different independent variables, we have plenty of leeway in making these analyses. It is odd, somehow, that the loss of interest in reactive inhibition occurred almost simultaneously with the appearance of new types of analyses. Reference is made to the work reviewed by Irion on the particular aspects of the performance which was changed by reactive inhibition, and by the work on bilateral transfer of reactive-inhibition effects. Such studies, it might seem, were the opening wedges into a new attack on the problems. History has not supported these expectations. Why?

We must first realize that the rapid decrease in the work on reactive inhibition did not occur because the characterization of the hypothetical process was complete. Far from it. We do not know the rate of buildup with any precision, particularly with regard to task variables with which the rate of buildup may be associated. There have been few studies in which transfer of the effects of reactive inhibition from one task to another was studied as a function of the task similarities. The characteristics of the dissipation have not been clearly determined. Is the rate constant regardless of the amount? Is it proportional to the amount?

Another possible reason for the loss of interest in reactive inhibition is that its presence in a wide variety of learning situations became doubtful. When reminiscence in verbal learning was a replicable phenomenon, temporary inhibition interpretations were applied in much the same way as they were subsequently applied to reminiscence in motor learning. Although we cannot rule out the possibility of a temporary work inhibition in verbal learning, it has been impossible to develop a situation from which it can be consistently inferred. It will be remembered that Hull introduced the notion of reactive inhibition primarily to handle extinction and spontaneous recovery. Contemporary theorists in animal learning make little use of the notion in explaining extinction and the work on spontaneous recovery has lagged so much that it is difficult to know in detail just what has to be explained. The general loss of interest in Hull's theory as a whole may in part be responsible for the contingent loss of interest in reactive inhibition. Selective learning studies have not been able to demonstrate phenomena for which reactive inhibition might be an appropriate explanatory concept. Thus it seems that there is a fairly restricted range of tasks which will produce the phenomena with which reactive inhibition is associated.

We may note that it is generally assumed (because the data demand it, as Irion has pointed out in Chapter 1) that reactive inhibition is only a performance variable—it does not influence learning. Given this situation, investigators whose interest lies in learning may have deliberately chosen skill tasks and practice schedules to avoid the contamination of their scores by reactive inhibition. However, the very facts which led to the conclusion that reactive inhibition influences only performance must at the same time be telling us something of considerable importance about learning. When a subject working under massed practice makes many fewer correct responses or makes less accurate responses than does a subject working under distributed practice and yet still learns as much or as well, we are presented with a fact of high relevance for learning theory.

Finally we note that the rapidly changing requirements of man's duties in various military roles may be responsible in part for the loss of financial support for studies in which tasks requiring rapid repetitive movements are used. Whatever the cause for the decline in the research on reactive inhibition, the fact remains that an important aspect of behavior still poses many problems to be solved experimentally, problems which are relevant to learning as well as to work-type behavior. Refined characterization of reactive inhibition seems doubly important in view of the fact that individual differences in the development of reactive inhibition are being related to personality characteristics, to vigilance behavior, and to elementary academic skills (e.g., Lynn, 1960; Otto & Fredericks, 1963).

## Retention

We now turn to a somewhat more substantive matter than those considered thus far. For many years, in one form or another, the question has arisen concerning the relative rates of forgetting of a motor skill and of a verbal list. It is probably not incorrect to say that a casual survey of the evidence on the matter would lead to the conclusion that verbal tasks are forgotten more rapidly than motor-skill tasks. However, as E. A. Bilodeau (Chapter 7) has pointed out, this difference, even if real, is probably of less magnitude than often believed. This is in part due to the fact that contemporary estimates of the rate of forgetting of verbal materials learned by naïve subjects are appreciably less than those given in the older literature, and in part due to the fact that Bilodeau interprets his data on retention of a positioning movement as showing more rapid forgetting than is commonly associated with skill tasks.

Adequate comparisons of the retention of verbal tasks and skill tasks are beset by many difficulties, some of which may be responsible for the apparent better retention of motor skills. (In discussing these difficulties there is some overlap with Irion's comments in Chapter 1.) Insofar as degree of learning influences acquisition, and insofar as the nature of our measurements do not allow an equation of degree of learning for different kinds of tasks, comparisons of retention are at once suspect. The response requirements for verbal and skills tasks often differ. High skill on a pursuit task may mean keeping the stylus on the center of the target; on a retention test performance may deviate considerably from this and still be counted as on target. In the recall of a verbal list, an item is usually counted as correct or incorrect. Thus, even though a subject may produce two of the three letters of a nonsense syllable correctly it is not counted as a correct response. In fact, the subject may not even attempt the response if he can only remember two of the three letters. It is quite possible that if we probed for recall of a verbal task, forcing responses when necessary, and scoring these responses even if not entirely correct, our estimates of forgetting of the verbal task would be appreciably reduced. Even if the material consists of words, such scoring could be accomplished in terms of correspondence between the number of syllables of the correct response and the given response, corresponding letters or syllables, and similarities in sounds.

In learning a skill task which is to be later tested for retention, it is common to give distributed practice in learning; the purpose is to eliminate the reminiscence component which may be present if acquisition is given under massed practice. In a number of studies, the practice period has extended over days. If a verbal task is given such forms of practice, at least when there is heavy proactive interference for the list to be recalled, forgetting is markedly retarded (Keppel, 1964). Adequate comparisons of the retention of verbal and motor skills must somehow equalize the tasks on this variable. Some investigators in studying the retention of motor skills (e.g., Fleishman & Parker, 1962) have eliminated the initial scores on a retention test on the grounds that these are depressed by warmup decrement. This assumption may or may not be correct; what is called warmup decrement may be called forgetting in a verbal-learning experiment. We obviously need to use warmup tasks before recall which will not interfere with the task to be recalled and still extract the warmup factor if there is one. But the results of attempts to meet this demand have been quite inconsistent (Adams, 1961).

Why should we be concerned about the relative retention of verbal skills and motor skills? The answer obviously lies in our desire to maintain continuity in our interpretation of behavior. We will probably avoid as long as possible the conclusion that forgetting of verbal skills and motor skills are based upon different processes or that the skills represent two different systems with little overlap. One long-held interpretation of forgetting of verbal material rests heavily on interference from conflicting associations. Modern refinements of the theory have used retroactive and proactive inhibition as studied in the laboratory as the key to overall forgetting. Thus, when a subject learns a single task and attempts to recall it, say, after one week, the retention will be interfered with by conflicting associations acquired throughout lifetime prior to the learning of the task to be recalled (proactive inhibition) and it will be interfered with by conflicting associations acquired during the one-week retention interval (retroactive inhibition). Certain refinements to the theory provide in some detail the mechanisms by which this interference is assumed to be governed (e.g., Underwood & Postman, 1960). Since relatively short retention intervals are normally used in our experimental studies, it becomes apparent that sources of proactive interference are primary. Essentially, we view the entire verbal repertoire which the subject brings to the experimental situation as being the source of the interference at recall. It has become necessary to map this associative network if we are to make predictions about differential rates of forgetting for different types of material. Thus, a task which requires the acquisition of associations which are directly contradictory to strong pre-experimental associations should suffer greater forgetting than one which is not.

Several tests of the role of pre-experimental associations in forgetting have been disappointing. However, certain recent experiments give some backing to the theory (see later part of Chapter 7, and Blick, 1965; Saltz, 1965). If the theory, at least in its general outlines, should subsequently prove appropriate (and most of us have no intention of abandoning it as yet), then in developing a relationship between the forgetting of verbal skills and motor skills, we should, perhaps, attempt to apply the theory to the motor skills. This would mean the mapping of the "associative network" of motor skills which the subject brings to an experimental situation. Just what this may mean operationally is difficult to say. However, one attempt (Melton, 1964b) has been made along these lines by assuming that the population stereotypes (S-R compatibility) represent a potential proactive source of interference if the subject is trained in a manner that is incompatible with the stereotype.

If we seriously set about to attempt to compare the retention of verbal and motor skills, we may wish to devise tasks which embody certain characteristics of both verbal and skill learning. Bilodeau, in his work with the retention of simple positioning movements, may have given us a lead of great value. Suppose we used, say, 10 discrete stimuli (analogous to stimuli in a paired-associate list) and trained the subject to make a movement of a particular distance to each. Each movement will have a limited range of tolerance for being counted correct. On a retention test we could not only measure the retention of the particular associations but also the precision with which the responses are executed, and these two factors may be measured independently. Proactive and retroactive inhibition could also be studied by making either an A-B, A-C paradigm or an A-B, A-B jumbled paradigm. In a somewhat similar fashion, a serial list could be simulated by eliminating the stimuli and requiring only that the subject learn a particular order of a series of positioning movements.

Finally we turn to certain issues raised by Bilodeau's studies (Chapter 7) on the retention of a simple positioning movement. Bahrick, in his discussion of Chapter 7, has raised serious objections to the studies, primarily on the grounds that nothing or little is known about the learning which occurred prior to the introduction of the retention interval, and knowledge of learning is considered essential before retention measurements can be given meaning. Bilodeau does not believe that a mean retention measure is a very good one, and has proposed the use of two other measures to index forgetting, namely, the correlation coefficient and the variance. More particularly, it is assumed that a decrease in correlations between learning measures and retention measures will be indicative of forgetting, while an increase in variance over time will also be indicative of forgetting. The intent of the present discussion is to examine the implications and relationships of these two measures for the study of forgetting. Some background is essential.

If we give a group of subjects a constant number of training trials on a verbal list and measure retention after 24 hours, the correlations will be quite high between the level of performance attained on the last acquisition trial and the recall trial. In three different verbal tasks, these correlations ranged from .74 to .81 (Underwood, 1964, Table 2). We have interpreted such relationships to mean that differences among subjects in level of learning attained are far more powerful determinants of amount forgotten than are individual differences in rate of forgetting. Similar data are available from the study of a complex motor skill. In a study by Fleishman and Parker (1962), which has also been discussed by Jones in Chapter 3, retention measures were made on various groups at intervals up to 2 years following a constant number of trials on the task. For seven different groups the correlations between performance scores at the termination of learning and scores on the retention test varied from .80 to .98. There was no clear change in these correlations as the retention interval increased. The magnitudes of these correlations are probably "inflated" somewhat by the fact that a relearning measure was used and the first 3 minutes of practice during relearning eliminated from the scores (to remove warmup effects). Thus the retention measures are not recall measures and probably include new learning. Nevertheless, the magnitude of these relationships corresponds well to those found in verbal-learning studies.

We may now examine the lever-positioning situation and see what may be expected to happen to variances and correlations and what implications these measures have for understanding forgetting. Let us assume we are going to teach subjects to move the lever 15 inches, and we are going to give all subjects 10 acquisition trials with knowledge of performance after each trial. On the first trial we tell the subject the general nature of the task and then ask him to guess on the first trial, giving an instruction such as, "Move the lever a distance that seems natural to you." After this trial, we tell him how much he was "off," and then give another trial, and so on, until the 10 trials are completed. Retention is measured by asking the subject to produce the movement he had learned during the training session.

We may now examine three possible indicants of forgetting (means, variances, correlations), given certain outcomes. We would first, of course, need some measure of original learning. Just what this should be is difficult to say without actually working in the situation, so let us assume that we can obtain a measure of learning for each subject in such a manner that we have some individual differences in the degree of learning. Let us assume further that we are going to view the results in terms of an interference theory, the interference being produced by the "natural" response bias of the subject in this situation, a bias which we have attempted to measure on the first trial.

(We in fact might find it necessary to give several "first" trials to get a more reliable measure of the bias.)

As one possible outcome, assume that the mean response on the first trial is 10 inches, indicating some population bias to undershoot. Assume further that one subject in this distribution moved the lever 5 inches on the first trial, a second subject moved it exactly 15 inches (he was on target the first trial). We may expect that the first subject will have more difficulty learning this task than the second and we would properly assess his degree of learning as lower at the end of the 10 acquisition trials. If forgetting occurs, the scores of the first subject should move back toward the 5-inch point, the scores of the second subject should remain essentially at 15. Relative to the performance on the last acquisition trial there would indeed be an increase in the variance as the retention interval increased, and complete forgetting would be indicated when the variance reached the same level present on the first acquisition trial. At the same time, however, the mean performance should shift downward in the direction of the mean setting measured on the first trial. Either the mean or the variance would be indicative of the same processes of forgetting. What about the correlations? Here we have to make a working assumption about the measures of degree of learning relative to the initial biases. The subject who set the position at 5 inches will be assumed to have a larger negative constant error than the one who set it initially at 15 inches. In a manner of speaking, the initial bias interferes with the learning; such interference would be most apparent if we gave no knowledge of the magnitude of the error on each trial—only the direction. Given this situation, the correlations among the response magnitude on the first trial, the degree of learning achieved in 10 trials, and the score on the recall trial should all remain high. In fact, on statistical grounds alone, one might expect an increase in the magnitude of the correlation between level of learning and the retention measure as the retention interval increases because of an increase in the variance of the individual scores. Thus, it would seem there is a no *a priori* reason to believe that a decrease in the correlations should occur as the retention interval increases. In fact, it would seem that the investigator would be hard pressed to explain such a decrease in correlation by any extant concepts. To explain it by an interference theory would require something like "random interference." But of course, interference may not be appropriate as an explanation for the forgetting of motor skills.

What would it mean if our two hypothetical subjects, clearly differentiated on the initial trial as well as on original learning as a whole, moved closer and closer together in their retention scores as the retention interval increased (indicating a decrease in the variance as well as a decrease in the correlation). That it seems unreasonable that this could happen does not, of course, mean that it could not happen. However, in view of previous findings it is most reasonable to expect that there will be an increase in variance with time but that the correlations between original level of learning and retention will remain relatively stable over time.

We next consider another possible outcome of our hypothetical experiment. Assume that on the initial trial the mean of the group is 15 inches; the average positioning movement is right on target, indicating no group bias for undershooting or overshooting, and the distribution of scores is symmetrical. Obviously, with 10 training trials the range of the distribution will be reduced. If forgetting occurs here there is no reason to expect a change in the mean scores on the retention test and, as Bilodeau has pointed out, if we look only at the means we would conclude that there has been no forgetting. But, in accordance with the previous discussion, we would expect an increase in variance; those who had an undershooting tendency should "regress" toward undershooting, those who had an overshooting tendency should regress toward overshooting. Again, however, there is no reason to expect the correlations to decrease. To have a decrease in correlation is possible, we repeat, but such a measure of forgetting means that as the retention interval increases the performance of the subjects during original training becomes less and less related to the retention scores, and if one cannot predict the retention of a subject on the basis of what we know about his history with a given task then attempting to understand the causes of forgetting becomes a somewhat hopeless affair.

Our general conclusion is that an increase in variance as a measure of forgetting can be a very meaningful measure in certain situations. However, we will expect this increase to be accompanied by substantial correlations between original learning and retention and those correlations will not decrease substantially with time. If they do, and if the decrease cannot be attributed to an artifact, such as many subjects falling to a zero level of performance, we are faced with a very difficult interpretative problem. On the basis of previous evidence it seems unlikely that we will have to face the problem.

## Some Final Observations

Research in psychology today is often said to be highly specialized. Sometimes this evaluation carries a negative component and we are said to be narrow. The investigators working in the field of motor-skills learning are probably less narrow than any other group; they essentially remain general psychologists because the nature of their research demands it. This demand may in part have been produced by the fact that many of their problems arose originally as applied problems—the man-machine relationship as it occurs in military and industrial settings. Learning to fly a modern airplane presents most of the problems with which much of general psychology concerns itself in abstracted forms.

Whatever the reason, the fact remains that these investigators cover a broad range of problems. They are psychophysicists; they study sensation and perception; they are factor analysts; some are cyberneticists. They study simple associative learning, probability learning, concept learning, as well as their specialty, skill learning. They study the effect of emotional states on skill learning; they study vigilance, attention, and reaction time. They must necessarily use concepts borrowed from other fields, but they have some which are primarily their property (e.g., *feedback*), and some which are not only unique to the field but also delightfully descriptive (e.g., *dither*). They have great technical skills, a source of embarrassment to those of us who are still trying to understand the wheel.

Regardless of the particular evaluation we may place on individual contributions to this book, one conclusion is inescapable: motor-skills learning is a broad, viable field—it is here to stay.

## REFERENCES

Adams, J. A. The second facet of forgetting: a review of warm-up decrement. *Psychol. Bull.*, 1961, **58**, 257–273.

Asch, S. E., & Ebenholtz, S. M. The principle of associative symmetry. *Proc. Amer. phil. Soc.*, 1962, **106**, 135–163.

Bilodeau, E. A. The relationship between a relatively complex motor skill and its components. *Amer. J. Psychol.*, 1957, **70**, 49–55. (a)

Bilodeau, E. A. Patterns of internal consistency in a multipart skilled performance. *Amer. J. Psychol.*, 1957, **70**, 550–559. (b)

Blick, K. A. Cultural primaries as a source of interference in short-term verbal retention. *J. exp. Psychol.*, 1965, **69**, 246–250.

Cohen, J. C., & Musgrave, B. S. Effect of meaningfulness on cue selection in verbal paired-associate learning. *J. exp. Psychol.*, 1964, **68**, 284–291.

Deese, J. The associative structure of some common English adjectives. *J. verb. Learn. verb. Behav.*, 1964, **3**, 347–357.

Fleishman, E. A., & Parker, J. F., Jr. Factors in the retention and relearning of perceptual motor skill. *J. exp. Psychol.*, 1962, **64**, 215–226.

Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.

Jenkins, J. J. Mediated associations: paradigms and situations. In C. N. Cofer and B. S. Musgrave (Eds.) *Verbal behavior and learning: problems and processes.* New York: McGraw-Hill, 1963, Pp. 210–244.

Keppel, G. Facilitation in short- and long-term retention of paired associates following distributed practice in learning. *J. verb. Learn. verb. Behav.*, 1964, **3**, 91–111.

Lynn, R. Individual differences in introversion – extroversion, reactive inhibition, and reading attainment. *J. educ. Psychol.*, 1960, **51**, 318–321.

Melton, A. W. Implications of short-term memory for a general theory of memory. *J. verb. Learn. verb. Behav.*, 1963, **2**, 1–21.

Melton, A. W. (Ed.) *Categories of human learning*. New York: Academic Press, 1964. (a)

Melton, A. W. Retention of tracking skills. Final Rep. Contract No. DA-49-007-MD-1020, U. S. Army Med. Res. and Develpm. Command and Univer. of Michigan (1964). (b)

Melton, A. W., & Irwin, J. McQ. The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer. J. Psychol.*, 1940, **53**, 173–203.

Miller, G. A. Some preliminaries to psycholinguistics. *Amer. Psychol.*, 1965, **20**, 15–20.

Otto, W., & Fredericks, R. C. Relationship of reactive inhibition and reading skill attainment. *J. educ. Psychol.*, 1963, **54**, 227–230.

Restle, F. Sources of difficulty in learning paired associates. In R. C. Atkinson (Ed.) *Studies in mathematical psychology*. Stanford, Calif.: Stanford Univ. Press, 1964.

Saltz, E. Spontaneous recovery of letter-sequence habits. *J. exp. Psychol.*, 1965, **69**, 304–307.

Underwood, B. J. Degree of learning and the measurement of forgetting. *J. verb. Learn. verb. Behav.*, 1964, **3**, 112–129.

Underwood, B. J., & Postman, L. Extraexperimental sources of interference in forgetting. *Psychol. Rev.*, 1960, **67**, 73–95.

Underwood, B. J., Ekstrand, B. R., & Keppel, G. An analysis of intralist similarity in verbal learning with experiments on conceptual similarity. *J. verb. Learn. verb. Behav.*, 1965 (in press).

Waters, R. H. The law of acquaintance. *J. exp. Psychol.*, 1939, **24**, 180–191.

Young, R. K. Tests of three hypotheses about the effective stimulus in serial learning. *J. exp. Psychol.*, 1962, **63**, 307–313.

# Author Index

(Numbers in *italics* refer to pages on which the complete references are listed.)

A

Abbey, D. S., 271, *292*
Abernethy, E. M., 17, *41*
Adams, J.A., 9, 20, 21, 24, 28, 30, 33, 34, 35, 36, *37*, *43*, *44*, 53, 55, 76, 83, 86, 87, *90*, *93*, 112, 115, 118, 119, *144*, 159, *165*, 172, 177, 179, 180, 183, 184, 185, 186, 187, 188, 189, 190, 192, 193, 194, 195, *197*, *198*, 203, 207, 208, 209, 211, *211*, *212*, 236, *240*, *242*, 251, *253*, 257, 277, *289*, 316, 329, *347*, *349*, 392, 401, 402, 403, *404*, *409*, 470, *478*, 510, *515*
Alcock, W. T., 31, *44*, 51, 53, 67, 78, 82, 83, 84, 86, 87, 89, *94*, *95*, 255, 266, 268, *294*
Allard, M., 263, 265, 285, *292*
Alluisi, E. A., 209, 210, *211*, *212*, 258, 271, 278, *289*, *291*
Ammons, C. H., 27, 28, *38*, 280, *289*, 316, *347*
Ammons, R. B., 5, 14, 19, 20, 21, 22, 23, 27, 28, 30, *38*, 237, *240*, 280, *289*, 316, *347*, 361, 369, 402, 403, *404*
Anderson, N. H., 54, *90*, 274, *289*
Andreas, B. G., 27, *38*, 86, *90*, 280, *289*
Annett, J., 11, 23, *38*, 237, *240*, 269, 272, 274, 275, 277, *289*, 300, *312*
Ansell, S. D., 431, 438, 439, 447, *481*
Archer, E. J., 22, 23, 33, *38*, *40*, 223, 236, *240*, 274, 277, *289*, 369, *404*
Arnoult, M. D., 176, *197*, 227, *240*
Asch, S., 219, *240*, 493, *515*
Atkinson, C. J., 438, *479*
Atkinson, J. W., 101, *107*
Atkinson, R. C., 106, *107*, 271, *289*
Attneave, F., 9, *38*
Axelrod, H. S., 88, *90*

B

Bahrick, H. P., 9, 30, *38*, 184, *197*, 256, 269, 270, 275, 277, 285, 288, *289*, *291*, *292*, *294*, 317, 318, 319, 335, *347*, 351, 356, 357, *359*, *360*, 369, 375, 383, 399, 403, *404*, *405*, 412, 419, *423*, *424*
Bahrick, P. O., 318, *347*
Bakan, P., 177, *197*
Baker, B. L., 87, *95*
Baker, C. H., 176, 188, *198*, 207, 210, *211*, *212*, 380, *404*
Baker, K. E., 25, 27, *38*, *42*, 217, 222, 235, *241*
Banks, H. P., 76, *95*
Barch, A. M., 26, 28, *38*, 53, 75, *90*, 236, *240*, 246, *253*, 288, *289*
Barnes, J. M., 332, *347*
Baron, M. R., 86, *91*
Bartlett, S. C., 207, 208, *212*
Battig, W. F., 28, 32, *38*, 215, 218, 219, 220, 223, 224, 227, 231, 233, 234, 235, 237, *240*, *241*, *242*, 262, 268, 273, 276, 280, 288, *289*, *291*, 306, *312*, 316, *347*, 374, 375, *404*
Bavelas, A., 255, *295*
Bechtoldt, H. P., 25, 36, *39*, *43*, 159, *165*

Becker, P. W., 265, 269, *289*
Begg, T. L., 258, *296*
Beinert, R. L., 207, 208, *212*
Bekey, G. A., 398, *404*
Bell, H. M., 13, 27, *39*
Ben-Avi, A. H., 143, *144*
Bennett, C. A., 288, *289*, 388, *404*
Bennett, W. F., 383, *404*
Berkshire, J. R., 138, *144*, *146*
Besch, N. F., 222, *240*
Besnard, G. G., 53, 77, *90*
Best, C. H., 203, *212*
Bilger, R. C., 438, *479*
Bilodeau, E. A., 3, 22, 26, 27, 28, 29, 30, 31, 32, *39*, *40*, *42*, *45*, 53, 55, 78, 81, 83, 86, *90*, 113, 125, 126, 133, 134, 135, *144*, 236, *240*, 246, 247, 251, 252, 253, 255, 256, 259, 262, 263, 264, 265, 266, 267, 269, 270, 271, 272, 273, 275, 276, 277, 278, 279, 281, 282, 283, 284, 285, 286, 287, 288, *289*, *290*, *291*, *292*, *295*, 303, 308, 310, *312*, 316, 320, 321, 322, 323, 324, 325, 326, 327, 328, 330, 331, 332, 334, 335, 337, 338, 339, 340, 342, 343, 344, 345, 346, *347*, *348*, *349*, 354, 356, 358, *359*, *360*, 368, 402, *404*, 470, *478*, 503, *515*
Bilodeau, I. McD., 3, 9, 26, 27, 30, 31, 33, *39*, *40*, *45*, 55, 78, 81, 83, 86, *91*, 236, *240*, 247, 251, *253*, 256, 258, 259, 262, 263, 265, 266, 268, 270, 271, 272, 273, 277, 279, 288, *290*, *291*, *295*, 303, 308, *312*, 316, 320, 321, 334, *347*, *350*, 402, *404*, 470, *478*
Birch, D., 53, 76, 78, *91*, 104, *107*
Birch, H. G., 129, *146*
Birmingham, H. P., 31, *42*, 261, 280, 288, *291*, *292*, *296*, 364, 372, 373, 376, *404*, *405*, *406*, *409*, 413, 414, 415, *423*, *424*
Black, J. W., 430, *478*
Blick, K. A., 246, 252, *253*, 271, *291*, 332, 338, 339, 340, 341, 342, 343, 344, *347*, *348*, *349*, 358, *360*, 511, *515*
Bloch, E., 28, *38*, 316, *347*
Boblitt, W. E., 268, *291*
Bojar, S., 459, 460, *481*
Borgatta, E. F., 109, *144*
Boring, E. G., 170, *198*, 203, *212*
Boswell, J. J., 328, *348*
Boulter, L. R., 172, 188, 189, *197*, *198*, 208, *211*
Bourne, L. E., Jr., 22, *40*, 236, *240*, 255, 263, 265, 267, 273, *291*, 303, 304, 305, 307, 310, 311, *312*
Bousfield, W. A., 219, *240*, 316, 336, *348*
Bowen, J. H., 374, *404*
Bower, G. H., 54, 71, *91*
Brackbill, Y., 268, *291*
Brackenridge, C. W., 88, *91*
Brackett, H. R., 224, 231, *240*, 268, *291*, 306, *312*, 316, *349*
Bray, C. W., 14, *40*
Bricker, P. D., 258, *294*
Bridgman, M., 460, *481*
Briggs, G. E., 9, 28, *38*, *44*, 178, *198*, 256, 274, 278, 288, *291*, *292*, *293*, *296*, 316, *349*, 351, *360*, 369, 370, 375, 376, 377, 383, 392, 398, 400, 401, 403, *404*, *405*, *407*, *410*, 412, 419, 421, 422, 423, *423*, *424*
Briggs, L. J., 53, 77, *90*, *93*
Britt, S. H., 14, *40*
Broadbent, D. E., 192, *198*, 207, 208, 209, 211, *212*, 316, *348*, 364, 365, *406*
Brody, D., 131, *144*
Brogden, W. J., 28, 32, *38*, 62, 65, 74, *91*, 262, 273, 276, 280, 288, *289*, 316, *347*, 368, 374, 375, *404*, *407*
Broussard, I. G., 32, *44*, 257, 284, *294*
Brown, G. G., 430, 447, *478*
Brown, I. D., 377, 378, *405*
Brown, J. S., 30, *40*, 86, *91*, 151, *165*, 256, 257, 262, *291*
Brown, L. K., 253, *253*
Brown, S. C., 219, 227, 231, *240*
Brown, W. L., 67, *95*
Bruce, R. W., 14, *40*, 222, *241*
Bryan, W. L., 1, 11, *40*
Bugelski, B. R.. 221, 224, *241*
Bunderson, C. V., 255, 263, 265, 267, *291*, 303, 305, 306, *312*

Bush, R. R., 53, 54, 71, *91*, 100, *107*
Buss, A. H., 308, *312*
Buss, E. H., 308, *312*
Buxton, C. E., 14, 15, 20, *40*, *45*

C

Cadwallader, T., 221, 224, *241*
Cajal, Ramon y, S., 459, *478*
Calvert, E. S., 370, *405*
Calvin, A. D., 76, *95*
Cambria, R., 463, *481*
Campbell, M., 15, *40*
Cantor, G. N., 79, *91*
Cantor, J. H., 84, *91*
Carterette, E. C., 271, *289*
Cason, H., 257, 281, *296*
Chambers, B., 275, *295*
Chambers, R. M., 398, *405*
Chambers, R. W., 81, 82, 89, *91*, *95*, 195, *197*, 258, *294*
Chambliss, D. J., 27, 28, *43*
Chang, J. J., 316. *350*
Chapanis, A., 371, 380, 384, *408*
Chase, R. A., 30, *40*, 262, *291*, 426, *478*
Chernikoff, R., 288, *291*, 372, 374, 377, 382, 385, 389, *404*, *405*, *409*, 413, 417, *424*
Chess, S., 129, *146*
Chiles, W. D., 210, *211*
Chinn, R. McC., 258, 271, 278, *291*
Clarke, F. R., 86, *91*
Cofer, C. N., 316, 337, *348*
Cohen, B. H., 219, 229, *241*
Cohen, J. C., 492, *515*
Cohen, L. B., 85, *93*
Coke, E. U., 337, *350*
Coleman, E. B., 252, *253*, 342, *348*
Conklin, J. E., 31, *40*, 261, 262, *292*, 369, 386, *405*, 418, *424*, 437, *478*
Connor, M. B., 382, *407*
Conrad, R., 337, *348*
Cook, J. O., 237, *241*
Cook, J. S., 371, 380, 384, *408*
Courtis, S. A., 59, 60, 61, *91*
Courts, F., 9, *40*
Cowen, E. L., 88, *90*
Crafts, L. W., 11, 14, *40*, 257, *292*
Craig, D. R., 365, *405*
Craik, K. J. W., 191, *198*, 364, 365, 366, 389, 390, *405*, *406*
Creamer, L. R., 184, 185, 187, 189, 195, *197*, *198*
Cronbach, L. J., 109, 111, *144*, 157, 165, *165*
Crowley, M. E., 24, 41, 248, 253
Cruzen, M., 411, *424*

D

Dallett, K. M., 221, *241*
David, E. E., Jr., 261, 288, *296*
Davis, H., 459, *478*
Davis, R., 192, *198*, 236, 237, *241*, 364, 365, 389, *406*
Davlin, D., 268, *291*
Day, R. H., 27, *41*, 235, *241*, 399, *406*
Dees, V., 255, 265, 280, *292*. *294*
Deese, J., 63, 72, *91*, 188, *198*, 207, 208, 210, 211, *212*, 218, *241*, 316, 337, 342. *347*, 493, *515*
Deininger, R. L., 280, *292*
Delafresnaye, J. F., 171, *198*
Denny, M. R., 79, *91*, 263, 265, 273, 285, *292*, 402, *406*
DeSoto, C. B., 28, *46*
Dey, M., 28, *38*, 316, *347*
Digman, J. M., 21, *41*
Dill, J. B., 53, 71, *96*
Ditchburn, R. W., 202, 207, *212*
Dodd, D., 303, 304, *312*
Dodge, R., 191, *198*
Dollard, J., 228, *242*
Donohoe, J. W., 221, *242*
Doré, L. R., 12, 13, 19, *41*
Dornbush, R. L., 219, *244*
Dulaney, D. E., 104, 105, *107*
Duncan, C. P., 23, 24, 25, 27, 28, *41*, *42*, 52, 75, 83, 84, *91*, *93*, 222, *243*, 247, 250, *253*
Duncan, D. B., 448, 456, *478*
Dunham, M. D., 79, *91*

E

Eason, R. G., 280, 288, *292*
Eaton, M. T., 269, 272, *292*
Ebbinghaus, H., 315, 317, *348*
Ebenholtz, S. M., 219, *240*, *241*, 493, *515*

Edgerton, H. A., 113, 122, *145*
Edwards, W., 54, *91*
Ekstrand, B. R., 305, *313*, 498, *516*
Elithorn, A., 192, *198*
Elkind, J. I., 369, 381, 397, *406*
Elliott, R., 210, *212*
Ellis, H., 227, *241*
Ellison, G. D., 151, 155, 156, *166*
Ellson, D. G., 26, *41*, 364, 365, 366, 367, 390, *406*
Elwell, J. L., 30, *41*, 269, *292*
Espenschade, A., 15, *41*
Estes, B. W., 310, *312*
Estes, W. K., 53, 54, 63, 71, 72, 82, *91*, *92*, 100, 101, 106, *107*, 173, *198*, 287, *292*
Ewert, P. H., 430, 447, *478*
Eysenck, S. B. J., 27, *41*

F

Fairbanks, G., 30, *41*, 426, 431, *479*
Fallon, D., 218, *241*
Farber, I. E., 88, *92*, 104, *107*
Farese, F. J., 3, *44*, 52, 75, 83, *92*, *94*, *95*
Farr, R. G., 28, *38*, 316, *347*
Feallock, J. B., 307, *312*
Feldman, S., 218, *241*
Fender, D. H., 460, *479*
Ferguson, G. A., 157, *165*
Finck, A., 280, *289*
Fishback, J., 24, *43*
Fitts, P. M., 9, *38*, *41*, 55, 76, 81, *91*, 162, *165*, 176, 178, *198*, 251, *253*, 256, 270, 274, 275, 277, 280, 288, *289*, *291*, *292*, *293*, 346, *348*, 351, *360*, 367, 369, 374, 376, 383, 390, 399, 403, *404*, *405*, *406*, *407*, *408*, 412, 419, 423, *423*, *424*, 426, *479*
Fleishman, E. A., 34, 35, 36, *41*, 55, *92*, 113, 115, 117, 120, 123, 124, 125, 132, 137, 144, *145*, 149, 151, 152, 153, 155, 156, 158, 159, 160, 161, 163, 164, *166*, *167*, 258, *292*, 316, 330, *348*, 510, 512, *515*
Foley, P. J., 271, *292*
Forbes, T. W., 379, *406*
Ford, A., 176, *198*, *200*
Foreman, S., 265, *293*
Foster, H., 24, 25, 27, *41*, *42*, 217, 222, *241*, 248, *253*
Fox, P. W., 252, *253*, 275, *292*, 338, 339, 340, 345, *348*, *349*, 358, *360*
Frankmann, J. P., 187, 188, 193, *198*, 203, 209, 211, *212*
Frankmann, R. W., 255, 256, 276, 286, *295*
Fredericks, R. C., 509, *516*
Freedman, D. G., 129, *145*
Freeman, F. N., 17, *41*
Freeman, P., 372, *409*
French, E. G., 258, *292*
Frick, J. W., 369, *406*
Frisbey, N., 402, *406*
Frost, G. G., 392, *406*
Fruchter, B., 160, 162, *166*
Frye, R. L., Jr., 84, *95*
Fuchs, A. H., 393, 399, *405*, *406*, 419, 420, 421, *424*
Fuchs, J. E., 52, 74, 78, 89, *92*, *94*, *95*, 258, *294*

G

Gagné, R. M., 24, 25, 27, *38*, *41*, *42*, 149, *167*, 217, 222, 235, *241*, 248, *253*, 278, 288, *292*, *294*
Gainer, C. A., 316, *349*
Galton, F., 128, *145*
Garmezy, N., 218, *242*
Garskof, B. E., 338, *349*
Garvey, W. D., 31, *42*, 261, 288, *292*, 366, 369, 376, 381, 392, 393, *406*, 415, 419, 420, 421, *424*
Gebhard, J. W., 378, *408*, 430, *479*
Geldard, F. A., 380, *406*
Gerjuoy, H., 53, *92*
Gerjuoy, I. R., 53, *92*
Gibbons, E. W., 438, *481*
Gibbs, C. B., 27, *42*
Gibson, E. J., 226, *241*, 491, *515*
Gifford, R. N., 182, *198*
Gilbert, R. W., 257, *292*
Gilden, L., 30, *40*
Golby, C. W., 237, *240*
Goldiamond, I., 438, *479*

Goldman, H. M., 23, *42*
Goldstein, D. A., 387, *406*
Goldstein, M., 278, *292*
Gompertz, B., 59, *92*
Goodenough, F. L., 15, *42*
Goodstein, L. D., 309, *313*
Gordon, N. B., 273, 280, 288, *292*
Goss, A. E., 220, 231, 235, *241*
Gottsdanker, R., 364, 365, 369, *406*
Gould, J., 30, *45*, 378, *406*, 449, *479*
Graham, D., 393, *406*
Graham, J. R., 207, 208, *212*
Granit, R., 459, *479*
Grant, D. A., 14, *40*, 52, 53, 54, 76, 78, 88, *90*, *92*, *93*, *95*
Gray, F. E., 366, 367, 390, *406*
Green, R. F., 27, *38*, 280, 288, *289*, *293*
Greenberg, R., 218, *241*, 250, *253*
Greene, E. B., 113, 145, 155, *167*
Greenfield, N., 235, *241*
Greenspoon, J., 265, *293*, *295*
Gregg, L. W., 375, *404*
Gregory, B. N., 220, 231, *241*
Gregory, R. L., 374, *409*
Grice, G. R., 23, *42*, 88, *92*, 261, *293*
Grindley, G. C., 30, *41*, 255, 265, 269, 280, *292*, *294*
Guilford, J. P., 132, *145*, 202, *212*
Guilfoyle, G., 30, *40*
Gustafson, L. M., 23, *42*
Guthrie, E. R., 48, 63, 66, 67, 71, *92*, 173, *198*
Guttman, L., 36, *42*, 122, *145*
Guttman, N., 430, *479*
Guy, D. E., 263, 267, *296*, 303, 304, 307, *312*, *313*

H

Haider, M., 209, *212*
Hake, H. W., 53, 54, *92*
Hall, E., 263, 265, 285, *292*
Hall, T. J., 209, 210, *211*
Hamilton, G. B., 48, *92*
Hamilton, H. C., 271, 272, *293*
Hammerton, M., 380, 382, 388, *407*
Hanley, C. N., 438, *481*
Hanlon, W. H., 388, 399, *408*
Harcum, E. R. 223, *241*
Harley, W., 280, *295*, 431, 447, *481*
Harlow, H. F., 48, 71, 72, 83, *92*
Harris, J. D., 379, *407*, *408*
Harris, S., 237, *241*
Harter, N., 1, 11, *40*
Hartman, B. O., 280, 288, *293*, 374, *407*
Harvey, S., 30, *40*
Hauty, G. T., 255, 258, 272, 278, *293*, *294*
Hawkes, G. R., 209, 210, *211*, *212*
Hebb, D. O., 172, *198*, 209, *212*
Heilizer, F., 88, *90*
Hein, A. V., 454, *479*
Heinemann, R. F. D., 366, 372, 385, *407*
Heinemann, S. H., 222, 223, *242*
Held, R., 454, *479*
Helmholtz, H. L. V. von., 446, 464, *479*
Helson, H., 383, 392, *407*
Hempel, W. E., Jr., 34, 35, 36, *41*, 55, *92*, 117, 120, 144, *145*, 151, 152, 155, 156, 159, 160, 161, *166*, *167*
Henry, C. E., 14, *40*
Henson, J. B., 288, *293*, 393, *406*
Herbert, M., 156, *167*
Heron, W. T., 28, *45*
Hertzig, M. E., 129, *146*
Hick, W. E., 192, *198*, 364, 389, *407*
Hicks, J. A., 15, *42*
Hilgard, E. R., 12, 13, 19, *41*, *42*, 48, 51, 62, *92*
Hill, H., 364, 365, 366, 390, *406*
Hille, B. J., 337, *348*
Hillix, W. A., 300, *312*
Hirsch, J., 380, *407*
Hoffeld, D. R., 368, *407*
Holding, D. H., 235, *241*, 271, *293*, 301, *312*, 459, *479*
Holland, J. G., 211, *212*, 288, *293*
Holland, M. A., 53, 71, *96*
Holmqvist, S. D., 377, 378, *405*
Holton, R. B., 53, 85, *96*
Holway, A. N., 177, *199*
Horenstein, B. R., 5, 20, *43*
Hornseth, J. P., 53, 54, *92*
Horton, D. L., 219, *241*
Houston, J. P., 222, *241*, 338, *349*
Houston, R. C., 113, *145*, 278, *293*
Hovland, C. I., 14, *42*, *45*, 83, *94*

Howell, D. C., 345, *348*
Howell, W. C., 178, *198*, 288, *293*, 375, 376, 377, 392, *405*, *407*
Howland, D., 382, *407*
Hubbard, W. R., 80, 81, *92*
Hubel, D. H., 450, *479*
Hufford, L. E., 28, *37*
Hull, C. L., 19, 30, 36, *42*, 55, 57, 58, 59, 62, 64, 65, 67, 72, 77, 79, 86, *92*, *93*, 172, *198*, 403, *407*, 446, 456, *479*
Humes, J. M., 172, *197*
Humphreys, L. G., 13, 15, *40*, *42*, 53, *93*
Humphries, M., 369, *407*
Hunt, D. P., 281, 288, *293*, 375, *407*
Hunter, W. S., 315, *349*
Hurvich, L. M., 177, *199*

**I**

Irion, A. L., 13, 20, 23, 27, 28, *42*, *43*, *44*, 72, 77, *93*, 176, *198*, 329, *349*
Irwin, J. McQ., 491, *516*
Isaacs, R. B., 268, *291*
Israel, H. E., 14, *45*

**J**

Jahnke, J. C., 27, 28, *42*
Jakobovits, L. A., 219, *242*
James, W., 456, *479*
Jasper, H. H., 171, 172, *198*, *199*
Jeffrey, W. E., 85, *93*
Jenkins, H. M., 207, *212*
Jenkins, J. J., 219, *241*, 253, *253*, 336, 338, *349*, *350*, 492, *515*
Jenkins, W. L., 382, *407*
Jenkins, W. O., 151, *165*
Johnson, L., 262, *294*
Johnson, R. B., 227, 228, 229, 233, *242*, 316, *349*
Johnston, W. A., 316, *349*
Jones, E. I., 27, 28, *42*
Jones, H. E., 79, 87, *93*
Jones, M. B., 29, 35, 36, *40*, *42*, 113, 120, 124, 126, 140, 141, 142, *144*, *145*, 159, 162, *167*, 320, 330, 331, 332, 337, 345, *348*, *349*
Jones, R., 30, *45*, 261, 280, *295*, 448, *481*
Jones, R. E., 176, *198*
Jones, R. E., Jr., 304, 306, *312*
Jones, T. A., 87, *95*
Jost, A., 9, *42*
Jung, J., 223, *242*
Justesen, D. R., 303, 304, *312*

**K**

Kadushin, I., 380, *407*
Kaess, W., 74, *93*
Kanfer, F. H., 265, *295*
Kanoti, G. A., 223, *242*
Kappauf, W. E., 460, *481*
Kappy, M. S., 268, *291*
Karlin, L., 272, 275, *293*, 392, 401, *407*
Katz, M. S., 375, *407*
Kaufman, H., 36, *46*
Kausler, D. N., 223, *242*
Kay, H., 11, 33, *38*, 237, *240*, 269, 274, 277, *289*, 300, *312*
Keller, B., 129, *145*
Kelley, C. R., 373, *407*
Kendler, H. H., 54, 72, *93*, 310, *312*
Kendler, T. S., 310, *312*
Kennedy, K. E., 220, 231, *241*
Kennedy, W. A., 78, *96*
Kent, G. W., 274, 277, *289*
Keppel, G., 72, *96*, 219, 223, *242*, 250, *254*, 307, *312*, 329, *349*, 498, 510, *516*
Kientzle, M. J., 16, 20, *43*
Kimble, G. A., 5, 20, 23, *43*, 63, 72, 86, *93*, 218, *242*, 392, 403, *407*
King, W. J., 178, *199*
Kinkade, R. G., 33, *43*, 277, 278, *293*, 400, 401, *407*
Kintsch, W., 268, *293*
Kjeldergaard, P. M., 219, *241*
Knowles, W. B., 288, *292*
Koehler, J., 431, 438, 439, 447, 461, *481*
Koerth, W., 3, *43*
Köhler, W., 24, *43*
Kohler, I., 454, *479*
Koonce, J. M., 27, 28, *43*
Koppenaal, R. J., 220, *240*
Korn, S., 129, *146*
Krebs, D. F., 438, *481*
Kremer, E. J., 151, *166*
Krendel, E. S., 394, 396, 397, 398, *408*
Kuhlman, C. K., 85, *93*

L

LaBerge, D. A., 101, *107*
Lair, C. V., 309, *313*
Lambert, W. E., 219, *242*
Larimer, G. S., 271, *293*
Larré, E. E., 266, 269, 272, *293*
Lavery, J. J., 255, 268, 271, 272, *292*, *293*
Lavik, J., 253, *253*
Lawrence, C., 192, *198*
Leavitt, H. J., 18, *43*
Lee, B. S., 431, 437, *479*
Leland, Q. A., 78, *92*
LeMay, M., 288, *291*, 377, 382, *405*
Leonard, J. A., 182, *199*
Levine, M., 416, 419, *424*, 437, *479*
Levy, C. M., 29, 31, *39*, *40*, 124, 126, *144*, 264, 275, 285, *290*, *291*, *293*, 316, 320, 322, 323, 324, 325, 326, 327, 328, 330, 331, 334, 335, 337, 346, *348*, 354, 356, *359*, *360*
Lewis, D., 24, 25, 26, *38*, *43*, 58, 61, 62, 64, 65, 83, 84, *93*, 236, *240*, *242*, 246, *253*, 316, *349*
Lichtenstein, M., 176, *198*
Licklider, J. C. R., 469, *479*
Lincoln, R. S., 27, *43*, 269, 270, 271, 272, 273, 275, *293*, 385, *407*, 437, *479*
Lindsley, D. B., 171, 173, *199*, 203, 207, 209, *212*
Lloyd, K. E., 307, *312*, 316, *349*
Lockard, R. B., 369, *406*
Locke, E. A., 156, *167*
Loess, H. B., 75, *93*
Long, J. B., 221, *242*
Lorge, I., 20, 30, 31, *43*, 261, 262, 265, 268, 269, 272, *293*
Luce, R. D., 202, *212*
Luh, C. W., 315, 319, *349*
Lund, M. W., 371, 380, 384, *408*
Lynn, R., 509, *516*
Lyon, D. O., 11, *43*
Lyon, V. W., 138, *144*

M

McAllister, D. E., 24, 25, *43*, 52, 83, 84, *93*, 235, 236, *242*, 316, *349*
McCloy, C. H., 15, *43*
McCollom, I. N., 15, *45*
McConnell, D., 288, *293*, 375, *407*
McCoy, D. F., 268, *293*
McCrary, J. W., 261, 280, *295*
McDermid, C., 448, *480*
McFann, H. H., 84, *93*
McFarland, R. A., 177, *199*
McGeoch, J. A., 7, 13, 17, *43*, *44*, 72, *93*, 315, *349*
McGinnis, E., 14, *44*
McGovern, J. B., 223, 238, *242*
McGrath, J. J., 211, *212*
McGuigan, F. J., 265, *293*
Mackworth, J. F., 176, *199*
Mackworth, N. H., 171, 176, *199*, 203, 207, 208, 209, 210, *212*, 271, *293*
McNemar, Q., 130, *145*
MacPherson, S. J., 265, *294*
McRuer, D., 393, 394, 396, 397, 398, *406*, *408*
Magni, F., 459, *480*
Magoun, H. W., 171, *199*
Malmo, R. B., 172, *199*, 210, *212*, *213*
Manatis, J. V., 224, *242*
Mandler, G., 53, 85, *93*, 222, 223, *242*, 302, *312*
Marion, R., 28, *38*, 316, *347*
Martin, E., 223, *242*
Marx, M. H., 86, *93*, 300, *312*
Mathias, R., 53, *92*
Meighan, T. W., 209, *212*
Melton, A. W., 3, 9, 15, 17, 23, *44*, 51, 52, *93*, 113, *145*, 150, *167*, 274, *294*, 316, 346, *349*, 351, *360*, 489, 491, 511, *516*
Melzack, R., 459, *480*
Mengelkoch, R. F., 316, *349*
Mergen, M., 431, 461, *481*
Merikle, P. M., 223, 224, 225, 231, *240*, *242*
Meryman, J. J., 86, *91*
Meyer, D. R., 9, *44*
Miles, G. H., 84, *93*
Miller, G. A., 207, *213*, 469, *479*, 500, *516*
Miller, L. B., 310, *312*
Miller, N. E., 228, *242*
Miller, R. B., 33, *44*, 275, *294*
Miller, R. E., 85, *94*
Milton, J. L., 176, *198*
Mitnick, L. L., 366, *406*, 419, 420, *424*
Montgomery, V. E., 84, *93*

Morales, M. F., 61, *96*
Morgan, C. L., 265, *294*
Morgan, C. T., 218, *242*, 371, 380, 384, *408*
Morgan, L. V., 265, *294*
Morgan, R. L., 27, *38*, 280, *289*
Morin, R. E., 52, 76, 78, *93*, 278, 285, *290*, *294*
Morrisett, L., Jr., 83, *94*
Moruzzi, G., 171, *199*, 459, *480*
Mortimer, R. G., 392, 401, *407*
Moss, S. M., 377, *408*
Mosteller, F., 53, 54, 71, *91*
Mote, F. A., 274, 277, *289*
Mowbray, G. H., 378, *408*
Mowrer, O. H., 364, *408*
Muckler, F. A., 288, *294*, 372, 388, 399, *408*, 413, *424*
Müller, G. E., 456, *480*
Muller, D. G., 227, *241*
Munn, N. L., 14, *44*
Munroe, P., 151, *167*
Murdock, B. B., Jr., 222, *243*, 307, *312*, 316, 331, 337, *349*
Murphy, J. E., 379, *407*
Murphy, J. V., 85, *94*
Mursch, G. A., 53, 77, *90*
Musgrave, B. S., 492, *515*
Mussina, C. M., 265, 269, *289*
Mysziewski, M., 431, 461, *481*

N

Nagel, E. H., 28, 32, *38*, 276, 280, 288, *289*, 316, *347*, 374, 375, *404*
Namikas, G. A., 33, *38*, 277, *289*
Naylor, J. C., 28, *44*, 316, *349*, 398, 399, *405*, *408*, 421, 422, *424*
Neimark, E. D., 74, *94*
Neumann, E., 28, *38*, 316, *347*
Newton, J. M., 387, *406*
Nicely, P. E., 207, *213*
Nicks, D. C., 151, *167*
Noble, C. E., 3, 31, 32, *44*, *46*, 49, 51, 52, 53, 55, 59, 60, 61, 63, 65, 67, 68, 70, 71, 73, 74, 75, 77, 78, 79, 80, 81, 82, 83, 84, 86, 87, 89, *91*, *92*, *94*, *95*, *96*, 219, *243*, 255, 257, 258, 261, 266, 268, 275, 284, *294*, *296*, 337, *349*
Noble, J. L., 51, 53, 67, 77, 78, 82, 84, 86, 87, *87*, 89, *94*, *95*, 261, 275, *294*
Noble, M. E., 256, 275, 277, 285, *289*, *292*, *294*, 351, 357, *359*, *360*, 367, 382, 390, *407*, *408*, 419, 423, *424*
Nodine, C. F., 220, 231, *241*
Nye, P., 460, *479*
Nystrom, C. O., 76, *95*

O

Obermayer, R. W., 288, *294*, 372, 388, 399, *408*, 413, *424*
Olson, M. W., 382, *407*
Ormond, E., 207, 208, *212*
Ornstein, G. N., 156, *166*
Osgood, C. E., 24, *44*, 221, 222, 224, *243*
Oswald, I., 173, *199*
Otto, W., 509, *516*
Overall, J. E., 67, *95*

P

Parker, J. F., Jr., 34, *44*, 113, 125, *145*, 151, 152, 155, 160, 162, 163, *167* 316, *348*, 510, 512, *515*
Partridge, J., 203, *212*
Paschal, F. C., 170, *199*
Patel, A. S., 88, *95*
Paul, N. T., 73, 83, *95*
Pavlov, I. P., 470, *480*
Payne, B., 13, *44*
Payne, R. B., 255, 258, 278, *293*, *294*
Pearl, B. E., 413, *424*
Pearson, R. G., 272, *294*
Pendleton, R. B., 273, *291*, 310, *312*
Perkins, F. T., 12, *46*
Perkins, M. J., 76, *95*
Perl, R. E., 112, 121, *145*
Persons, R. W., 265, 269, *289*
Peters, W., 364, *408*
Peterson, J., 63, *95*, 447, *480*
Peterson, J. K., 447, *480*
Peterson, L. R., 219, *243*, 308, *312*, 316, *349*
Peterson, M. J., 308, *312*, 316, *349*
Phillips, R. S., 430, *480*
Pikler, A. G., 379, *407*, *408*
Pilzecker, S., 456, *480*

Pollack, I., 262, *294*
Poole, G. B., 17, *46*
Porter, L. W., 222, *243*
Postman, L., 27, *46*, 218, 219, 223, 225, 232, 238, *243*, 249, *254*, 316, 337, 341, *349*, *350*, 510, *516*
Poulton, E. C., 9, *44*, 182, 183, 189, 192, *199*, 273, 277, 288, *294*, 362, 364, 365, 366, 367, 369, 371, 372, 374, 375, 378, 385, 389, 390, 391, *408*, *409*, 411, *424*
Prentice, W. C. H., 72, *95*
Pritchard, R. M., 460, *480*
Pronko, N. H., 454, *481*
Puff, C. R., 316, *348*
Pyle, W. H., 9, 14, *44*

R

Rapin, I., 30, *40*
Ratliff, F., 460, *480*
Regan, J. J., 377, *409*
Reitz, J. G., 66, *96*
Renshaw, S., 13, *44*
Restle, F., 53, 54, 71, 82, *95*, 498, *516*
Reynolds, B., 21, 23, 33, 34, 35, 36, *37*, *42*, *44*, 113, 116, 120, *145*, 159, *167*, 316, 330, *350*, 392, 401, 402, *409*
Reynolds, D., 192, *199*
Reynolds, W. F., 222, *240*
Rhule, W., 288, *294*, 447, 463, *480*
Rich, S., 160, 162, 163, *167*
Richardson, J., 229, *243*
Riggs, L. A., 460, *480*
Riley, D. A., 74, *95*
Rittenhouse, C. H., 278, *292*
Robel, D. P., 89, *95*, 258, *294*
Robinson, D. N., 288, *295*
Robinson, E. S., 10, 14, 28, *44*, *45*, 220, *243*
Rockway, M. R., 23, *45*, 386, 388, 399, *408*, *409*, 412, 417, *423*, *424*
Roehrig, W. C., 316, *350*
Rogers, O. E., 139, *146*
Rokeach, M., 263, 265, 285, *292*
Rosenbach, J. H., 281, *290*
Rosenberg, S., 256, 269, 286, 287, *294*
Rosenquist, H. S., 24, 33, *40*, *45*, 277, *291*
Ross, S., 258, *294*
Rossi, G. F., 171, *199*
Rothkopf, E. Z., 337, *350*
Rouanet, H., 256, 269, *294*
Rubin, G., 236, *243*
Ruch, T. C., 10, *45*
Rund, P. A., 415, *424*
Runquist, W. N., 218, 219, *243*
Russell, W. A., 336, 338, *349*, *350*
Ryan, F. J., 26, 31, *39*, *45*, 265, 267, 285, *290*, 295

S

Saltz, E., 511, *516*
Saltzman, I. J., 255, 265, 268, *295*, 304, *312*
Samuels, I., 171, *199*
Sanders, A. F., 177, 178, *199*
Sanderson, W. A., 227, *243*
Saufley, W. H., Jr., 9, *45*, 219, *242*, 258, *295*
Schild, M. E., 219, 231, *240*
Schlank, M., 454, *479*
Schlosberg, H., 18, *43*, 51, 72, *96*, 201, 202, 204, 208, *213*
Schmidt, R. E., 74, *91*
Schneider, R., 270, *289*, 383, *404*
Schoenberger, R. W., 138, *146*
Schulz, R. W., 218, 219, 223, 229, *243*, *244*, 298, 305, *313*
Schumsky, D. A., 270, 284, *290*, *295*, 308, *312*
Schuyler, W., 1, 12, *46*
Schwartz, M., 249, *254*
Schwarzbek, W. C., 13, *44*
Scott, T. H., 209, *213*
Searle, L. V., 364, *409*, 472, *480*
Sears. R. R., 14, *42*, *45*
Seashore, H., 255, *295*
Seashore, R. H., 15, *45*
Seibel, R., 53, 76, *95*, *96*
Seidenstein, S., 368, 374, *407*, *409*
Senders, J. W., 411, *424*
Senders, V. L., 178, *199*
Serio, F. R., 388, 399, *408*
Servos, G. H., 431, 438, 439, 447, *481*
Seward, J. P., 53, 71, *96*, 397, *409*
Seymour, W. D., 421, *424*
Sharpless, S., 172, *199*
Shelly, C., 275, *289*

Shelly, M. W., 28, *46*, 288, *293*, 375, *407*
Shepard, R. N., 316, *350*
Shock, N. W., 61, *96*
Shoup, G. W., 151, *166*
Sidman, M., 215, *243*
Sidorsky, R. C., 271, *289*
Sidowski, J. B., 53, *96*
Siegel, J., 288, *295*
Siipola, E. M., 14, *45*, 430, 447, *480*
Silver, R. J., 28, *45*
Simon, B. P., 280, *295*
Simon, C. W., 178, *198*, 288, *292*, 376, *406*
Simon, H. A., 74, *96*
Simon, J. R., 280, *295*, 413, *424*
Skinner, B. F., 82, *96*, 429, 457, 469, *480*
Slack, C. W., 365, *409*
Small, A. M., Jr., 375, *404*
Smart, R. C., 15, *42*
Smelkinson, N., 268, *291*
Smith, C. J., 459, *480*
Smith, K. U., 9, 27, 30, *43*, *45*, 236, 237, *241*, *243*, 261, 280, 288, *294*, *295*, 378, *406*, 413, *424*, 426, 431, 436, 437, 438, 439, 444, 445, 447, 448, 449, 451, 453, 457, 459, 460, 461, 463, 468, *479*, *480*, *481*
Smith, M. B., 13, *42*
Smith, M. F., 426, 447, 457, *481*
Smith, P. N., 25, *43*
Smith, R. P., 207, 208, 210, *211*, *213*
Smith, S., 280, *289*
Smith, W. M., 9, 30, *45*, 261, 280, *295*, 426, 431, 445, 447, 448, 451, 453, 460, *481*
Smode, A. F., 255, 275, 277, 278, *295*
Snoddy, G. S., 12, 13, *45*
Snyder, F. W., 454, *481*
Sokolov, Y. N., 464, 471, *481*
Solandt, O. M., 202, *212*
Solley, C. M., 272, *295*
Spear, N. E., 305, *313*
Spearman, C., 114, *146*
Spence, J. T., 258, *295*, 309, *312*
Spence, K. W., 51, 54, 55, 57, 58, 59, 63, 72, 76, 77, 79, 88, *92*, *93*, *96*, 101, *107*, 261, *295*
Spiker, C. C., 53, 76, 79, 85, *91*, *96*
Spong, P., 209, *212*
Spragg, S. D. S., 27, *38*, 280, *289*, 375, *407*
Standfast, S., 30, *40*
Stanley, R., 280, *295*, 431, 447, *481*
Starch, D., 9, 11, *45*
Stark, L., 461, *482*
Steffan, J., 463, *481*
Stein, L., 86, *91*
Stenson, H. H., 172, *197*
Stevens, S. S., 202, *213*
Stockbridge, H. G. W., 275, *295*
Stratton, G. M., 30, *45*, 430, 447, 454, *481*
Straughan, J. H., 53, *92*
Suddon, F. H., 255, 268, 272, *293*
Sulzer, J. L., 29, 31, *39*, 264, 271, 276, 285, *291*, *295*, 322, 326, 334, 335, *348*, 354, *360*
Suppes, P., 255, 256, 269, 276, 286, *295*
Sutton, S., 30, *40*
Swartz, W. F., 288, *294*, 372, *408*, 413, *424*
Sweeney, J. S., 31, *42*, 261, *292*, 376, *406*
Swift, E. J., 1, 12, *45*, *46*
Szafran, J., 288, *295*

**T**

Tackaberry, T. H., 286, *295*
Taub, H. A., 220, 231, *241*
Taubman, R. E., 255, *296*
Taylor, A., 31, *46*, 79, 81, 82, 94, *96*, 255, *296*
Taylor, F. V., 280, 288, *291*, *296*, 364, 369, 372, 373, 374, 381, 385, 389, 392, 393, *404*, *405*, *406*, *409*, 413, 414, 417, *423*, *424*
Telford, C. W., 191, *199*
Terrell, G., Jr., 78, *96*
Thomas, A., 129, *146*
Thomas, P., 151, *167*
Thompson, A. S., 231, *243*
Thompson, J. D., 52, *95*
Thorndike, E. L., 11, 14, 28, 30, 31, *43*, *46*, 48, *96*, 191, *200*, 261, 262, 265, 268, 269, 272, *293*, *296*, 429, *481*
Thornton, G. B., 381, *409*

Thurstone, L. L., 74, *96*
Tickner, A. H., 380, *407*
Tiffany, W. R., 438, *481*
Tipton, C. L., 415, *424*
Titchener, E. B., 170, *200*
Tolman, E. C., 67, *96*, 99, *107*
Trowbridge, M. H., 256, 281, *296*
Tsai, L. S., 316, *350*
Tulving, E., 316, 331, *350*
Tupper, R. W., 263, 267, *296*, 303, 307, *313*
Twedt, H. M., 222, 223, *243*

**U**

Underwood, B. J., 18, 27, *46*, 54, 72, 84, *91*, *96*, 173, *200*, 218, 219, 221, 222, 223, 225, 226, 229, 232, 238, *241*, *242*, *243*, *244*, 247, 250, *253*, *254*, 278, *296*, 298, 305, 307, *312*, *313*, 316, 329, 332, 336, 337, 341, *347*, *349*, *350*, 359, *360*, 498, 510, 512, *516*

**V**

Valentine, W. L., 113, 122, *145*
van Bergeijk, W. A., 261, 288, *296*
Vandenberg, V., 76, *91*
Vanderplas, J. M., 227, *243*
Vanderplas, J. N., 227, *243*
Van Sant, C., 364, 365, *406*
Vince, M. A., 364, 366, 389, 390, *409*, 423, *424*
Viteles, M. S., 113, *146*
Voeks, V. W., 66, 71, *96*
Voss, J. F., 28, *38*, 262, 273, 276, 288, 289, 316, *347*, 375, *404*, *409*

**W**

Wagner, J. E., 268, *291*
Walk, R. D., 73, *96*
Walker, E. S., 53, 77, *90*
Walker, L. C., 28, *46*
Wallace, J., 310, *313*
Walters, H. A., 89, *95*
Walton, D., 258, *296*
Wargo, L., 30, *45*, 261, 280, *295*, 448, *481*
Warkentin, J., 460, *481*
Warren, C. E., 366, 390, *408*
Warrick, M. J., 261, *296*, 376, *410*, 416, 418, *424*, 430, 437, *481*
Waters, L. K., 398, *405*, 421, *424*
Waters, R. H., 17, *46*, 66, *96*, 494, *516*
Watson, J. B., 170, *200*, 429, 457, *481*
Waugh, N. C., 174, *200*
Weaver, J., Jr., 402, *406*
Webber, C. E., 179, 180, *197*
Wehrkamp, R., 236, 237, *241*
Weinberg, D. R., 263, 267, *296*, 303, 307, *313*
Welford, A. T., 192, *200*, 390, *410*
Wenborne, A. A., 364, *408*
Wheeler, L., 365, *406*
Wheeler, R. H., 12, *46*
Wherry, R. J., Jr., 138, *146*
Whipple, G. M., 15, *46*
White, B. N., 76, *96*
White, C. T., 176, *198*, *200*
White, J., 271, *293*
Wiener, E. L., 383, *405*
Wiener, N., 426, *481*
Wiesel, T. N., 450, *479*
Wilkinson, R. T., 300, *313*
Williams, A. C., 274, 278, *296*, 401, *410*
Williams, J. G., 219, 220, 231, *240*
Williams, J. M., 219, 220, 231, *240*
Willig, L., 21, *38*
Wilson, D., 268, *291*
Wimer, R., 221, *244*
Winchester, R. A., 438, *481*
Winnick, W. A., 219, *244*
Witte, R. S., 54, *96*
Wolfe, J. B., 446, *482*
Wolfle, D., 30, *46*, 139, *146*
Woodhouse, M. C., 377, 378, *405*
Woodrow, H., 15, 16, 17, 25, 35, 36, *46*, 86, *96*, 115, 118, 119, 122, *146*, 184, *200*
Woodworth, R. S., 11, 14, *46*, 51, 72, *96*, 191, *200*, 201, 202, 204, 208, *213*, 423, *424*
Wooster, M., 430, 447, *482*
Wykoff, L. B., Jr., 53, *96*, 176, *200*
Wylie, R. C., 25, 27, *38*

**X**

Xhignesse, L. V., 185, 186, *197*

Y

Yarmolenko, A., 15, *46*
Yates, A. J., 437, *482*
Yelen, D. R., 219, *244*
Yerkes, R. M., 48, 49, *96*, *97*
Young, L. R., 461, *482*; *482*
Young, R. K., 218, 219, 221, *244*, 493, *516*
Yntema, D. B., 316, *350*

Z

Zanchetti, F., 171, *199*
Zavala, A., 156, *167*
Zeaman, D., 36, *46*, 74, *93*
Zeiler, M. D., 85, *97*
Zuercher, J. D., 177, *200*

# Subject Index

Abilities, 111, 131–132, 137, 147–167
Activity, generation of, 474
Adjusted learning procedures, 224
Analysis, component processes, 219–220, 223, 233, 235–238, 245, 486–488
  physical characteristics as basis, 236–238
  *see also* Composition rules, analysis and synthesis
Anticipation
  in complex tasks, 185–187
  in vigilance tasks, 187–190
  mechanisms, 184–185
  perceptual, 182–183
  receptor, 182
  temporal, 183–184
Apparatus and tasks
  finger positioning, 234–235
  lever positioning, 320
  multiple choice, *see* selective learning below
  multiple-choice recognition, *see* Recognition, multiple-choice
  multiple-response, 226
  paired associate, 219–220, 222–223, 227, 231–232, 238
  psychomotor tests, 153–155
  selective learning, 47–56
  serial, 219, 231, 493
  tracking, 361–389, 411–419
  verbal discrimination, 219–220, 231
Arousal, 171
  theory and vigilance, 209–210
Association(s)
  backward (R-S), 218–220, 223, 226, 493, 504
  bidirectional properties, 493
  context, 223, 493
  facilitation networks, 251–252
  forward (S-R), 223, 226
  frequency, 505
  in motor learning, 494–495
  in verbal learning, 492–493
  mediated, 226, 345, 492–493
  unit of analysis, 2–3, 492–495
  *see also* Retention, association
Associative
  interference, 230–231, 251–252
  matching, 225
  stage of learning, 223, 238, 251
Attention, 202–204, 497
  and distraction, 205
  and vigilance, 170–171, 202–209
  classical study of, 204–206
  concept of, 173, 196–197
  definition, 172–175
  determiners, 204
  divided, 205
  early history, 169–170
  five classical areas of research, 204–205
  free and controlled, 204
  mechanisms, 169–213
  reticular formation and, 171–172
  shifting and fluctuation, 204–205
  span, 205
  verbal learning and, 173–174
  *see also* Monitoring, Signal detection, Vigilance, Watchkeeping

Compatibility, *see* Congruence, stimulus-response
Component(s)
  of movement, 429, 433–436, 450–454
  of trials, 259, 262–263, 302–308, 321–323, 333–336
  *see also* Analysis, component processes
Composition rules, analysis and synthesis, 10–11, 25–26, 57–63, 99–101, 133–136, 158–163, 219–220, 223, 233, 235–238, 259, 262–263, 305–307, 320–323, 333–336, 346, 393–398, 414–416, 419–421, 429, 433–436, 450–454, 495–504
  *see also* Whole-part
Concept
  formation, 263, 265, 310
  identification, 302–304, 310–311
Congruence, stimulus-response, 228–229, 231
Contiguity, 505–506
Correct paired-associate procedure, 224
Correlation(s)
  in acquisition, 15–17, 20, 34–36, 112–124, 158–163, 330–332, 353–355, 357, 511–514
  in retention, 124–127, 163–164, 330–332, 344–345, 351–359, 512–514
  intratask, 134–136
  multiple, 137
  patterns, *see* Factors, Monotonic hierarchy, Superdiagonal forms
  superdiagonal, 114–117
  test trial, 118–120
Countertraining, 267–268, 285–286
Cybernetic(s), 425–488
  closed- versus open-loop systems, 426–427
  computer techniques, 431–434, 438–445, 460–461, 486–487
  experimental techniques, 430–433
  interpretation of transfer of training, 456–459, 475
  nature of
    perception, 459–464, 473–474
    reflex, 429–431, 487
  research areas, 430, 440
  systems, properties of, 426–427
  television techniques, 444–445, 448–449, 451–452, 465–467
  theory and learning theories, 469–478, 484–488
Differentiation
  Gibson's theory, 226
  stimulus predifferentiation, 227, 237
  stimulus-response, 226
Distribution of practice, 506–509
  and retention, 321
  in history of skill, 9–10, 12–14, 19–24
  in positioning, 265–266, 302
  in selective learning, 60, 81–82
  in tracking, 402–403
  in verbal paired-associate learning, 307
  present status, 506–509

Emergent properties, 500, 503–504
Engineering variables in tracking, 363
Extinction, experimental, 504, 508
Eye movement, 176–178, 459–463

Facilitation
  and interference, 215–254
  in retention, 315–347
  in selective learning, 83–86
  interpair, 227
  intertask, 218, 220–236, 249–252
  intratask, 83, 226–233
  motor-skills, 233–239, 247
  *see also* Interference, Transfer of training
Factors
  g, 118
  improvement, 118
  number, 132, 157
  oscillating, 122
  psychomotor, 132–133, 150–158
  *see also* Correlation
Familiarization, 219–220, 231
Feedback, 29–34, 77–81, 255–313, 425–488
  action, 275–276
  and retention, 320–323, 326, 332–335
  augmented, 260, 269, 274–278, 310, 400–401
  compared with
    knowledge of results, 255–256, 466, 469–470, 485

    reinforcement, 255–259, 308–310, 445–446, 466, 470–472, 484–485, 505–506
  contingencies, *see* Feedback, action, augmented, delayed, displaced, error, frequency, locus, mode, serial, transformations
  control
    in speech, 468–469, 476–477
    mechanisms, 427–429, 450–452, 459–460, 483–484
  definitions, 255–259, 270–271, 299
  delayed, 77–78, 262–269, 302–308, 333, 418, 436–446, 484–485
    and learning theories, 436–438, 445–446
    auditory, 440–442
    learning with, 436–444, 462–463, 484–485
    visual, 442–445
    *see also* Feedback, locus
  displaced, 431, 446–456, 483–484
    and transfer, 454–456
    breakdown functions, 428–429, 447–450, 463, 483
    differential effects of, 452–454
    learning with, 447–454, 462–463, 467
    *see also* Feedback, transformations
  dynamic versus static, 466–468
  error
    increased-reduced, 282–283
    rounding, scale-grain, 279–282, 288
  extra, *see* Feedback, augmented
  extrinsic, 269, 274, 299
  frequency, 79–81, 260, 269–274, 308–309
  functions of, 299–302
  increased error in, 282–283
  information, 255–313
  instrumental, 464–469
  intermittent, 273
  intrinsic, 269, 274, 299
  kinesthetic, 269–271
  learning, *see* Feedback, action, and serial
  locus, 77–78, 260–269, 302–308
    delay, 259, 261–268, 302–308, 320–323
    post-information feedback interval, 259, 262–267, 302–308
  mode, 305–307
  operational, 464–469
  procedures, 260, 279–280
  properties, 257–259, 299–302, 483–488
  reactive, 464–469
  reduced error in, 282–283
  relative frequency, 272–274
  reminder, 263, 265
  role in learning and retention, 77–81, 255, 257–259, 269, 277, 299–302, 320–323, 333–335, 487–488
  scale-grain, 279–282, 287–288, 310
  serial, 77–78, 261, 275–276
  sources in tool using, 464–466
  summary, 400–401, 403
  supplementary, *see* Feedback, augmented
  terminal, *see* Feedback, action, serial
  terms, 255–256
  theoretical properties, 77–81, 257–259, 299–302, 483–488
  transformations, 77–81, 260, 279–288, 308, 310–311, 321, 427, 470, 475–477, 483–486
    comparative effects, 440–444, 462–463
    instrumental, 464–465, 475
    numerical, 279–288, 308, 310–311, 321
    physiological, 434, 463–464
    symbolic, 465, 468–469, 476–477
    *see also* Feedback, action, augmented, delayed, displaced, error, frequency, locus, mode, serial
  verbal reinforcement combinations, 258, 300, 308–309, 320–323, 326–327, 333–337
  *see also* Knowledge of results, Reinforcement
Forgetting, 315–360
  atomistic and dynamic theories, 359
  *see also* Retention
Frequency, 505
Functionalism, 490, 495, 504
Generalization, 85–86, 504
Genetic components

  changes with practice, 130–131
  in performance, 128–131
  in psychomotor tasks, 130
Gradient
  of delay of information feedback, 260–261, 264–265
  of reinforcement, 260–261
Guided practice, 271, 301
Habit-family hierarchies, 501
History of skill, 1–46, 490–492
  age-maturation hypothesis, 12–13
  Ammons' theoretical analysis, 19, 21
  composition rules, 10–11, 25–26
  correlational studies, 15–17, 20, 34–36
  definition of skills field, 1–4, 489–490
  distribution of practice, 9–10, 12–14, 19–24, also: 27
  fatigue, 12–13, 19–22
  feedback, 29–34
  historical periods, 4–8
  Hull's theoretical analysis, 5, 19–22, 30
  individual differences, 15–17, 20, also: 34–36
  inhibition
    proactive and retroactive, 10–11, 14, 24–25
    reactive, 19–22
  knowledge of results, 29–34
  law of effect, 29–34
  learning curves, 12
  length-difficulty, 10
  productivity
    in skills field, 4
    in transfer, 215–217
    of individuals, 8
  psychometrics, 14–17, 34–36, also: 20
  reflex, 2–3
  reminiscence, 20–24, 27
    bilateral, 23–24, also: 9–10, 12–14, 19–24
  response
    availability, 3–4
    selection, 3–4, 47–54
  retention, 12, 27–29
    relative, of verbal and motor habits, 17–19
  $_{S}I_{R}$, 19–22
  Snoddy hypothesis, 13–14
  stimulation-maturation hypothesis, 12–14
  systematic viewpoints, 8–9
  transfer of training, 11–12, 14, 24–27, 215–221
    between tasks of different difficulty, 26–27
    bilateral, 14, also: 23–24
  warming-up, 18, 27–29
  whole-part, 10–11, 25–27
  work, 12–13, also: 19–22
Homunculus, 173
Human data processing, limits of, 190–196, 486–487
Hypothetical construct, 491, 507
Incongruence, *see* Congruence
Individual differences, 15–17, 20, 34–36, 109–167
  in retention, 345, 354–358, 512
  in selective learning, 86–89
Information
  analysis, 504
  feedback, 255–313
  *see also* Feedback, Knowledge of results, Reinforcement
Inheritance, *see* Genetic components
Inhibition
  conditioned, 392, 403
  proactive, 510–511
  reactive, 403, 506–509
  retroactive, 491, 510–511
  *see also* History of skill
Instrumental learning, 464–468
Interference, 215–254
  associative, 230–231, 251–252, 485, 499–500
  in retention, 315–347, 510–514
  in selective learning, 83–85
  interpair, 227
  intertask, 218, 220–226
  intratask, 226–236
    motor skills, 222, 233–236, 245–247
    proactive, 218, 250–251
    retroactive, 218, 245–246
    sources of, 230–231
  random, 513
  *see also* Facilitation, Transfer of training
Interpolated activity, 268–269, 271, 302,

305–308, 328–329, 345
  in selective learning, 83
Interresponse or intertrial interval
  *see* Distribution of practice
Knowledge of results, 125–127
  compared with feedback, 255–256, 466, 469–470, 485
  *see also* Feedback
Lag time, 506
Language, 500–501, 503
Learning
  all-or-none, 505
  degree of, 509, 513
  free, 493, 498, 502
  incremental, 505
  set (learning to learn), 216, 222, 249–250
  theories of, 48–49, 56–73, 484–488
  *see also* specific topics
Meaningfulness, 219
  effect on transfer, 223–226, 251
Mechanism(s)
  of motor responding, 169–200
  of response, 201
    mechanics of, 210–211
    psychological study of, 201–202
    vigilance as a, 203
Mediation, 219, 225
Memory, *see* Retention
Methods of psychology
  differential, 110–112, 165
  experimental, 109–110
Monitoring, 202–204, *see also* Attention, Signal detection, Vigilance, Watchkeeping
Monotonic hierarchy, 126–127, 331–332, 344–345
Motor, 3–4
  skills (*see* Facilitation, Interference, Transfer of training)
  tasks
    compared with other tasks, 1–4, 297–299
    compared with verbal, 53–54, 233–239, 245–251, 298–299, 333, 345–346, 362, 489–516
    positioning, 259, 262–273, 275–276, 280–287, 320–323, 335, 355
  tracking, 361–424
Multiple choice learning, *see* Selective learning
Neural action, synaptic versus neuronic, 429–430, 487
Neutrality (in transfer), 221–222
  one-trial learning, 219
Noise, 288, 375–376
Observing response
  and eye movements, 176–178
  and inferences from motor responses, 178–181
  and multidimensional tracking, 178–181
One-channel hypothesis
  and event uncertainty, 194–196
  and the expectancy view, 192–194
  definition of, 190–191
Pacing, 361–363
Paired-associate task, 299, 305–308, 337, *see also* Apparatus and tasks
Paradigm, *see* Transfer of training
Part-whole, *see* Whole-part and *see also* Analysis, component processes, Composition rules, analysis and synthesis
Perception, cybernetic nature of, 459–464, 473–474
Personnel selection
  Air Force, 137
  follow-up, 138–139
  Navy, 138–139
  validity, 136–138
Population stereotypes, 164–165, 346, 511
Positioning responses, 259, 262–273, 275–276, 280–287, 320–323, 333–335, 355, 423
  procedures, 259–260
  trial, 259, 262–263
Post-information feedback interval, *see* Feedback, locus
Precision in tracking, 362
Prediction in tracking, 369, 391–392, 397
Pretraining, *see* Transfer of training
Probability learning, 53–54, 257, 273–274, 285–287, 308, 311, 497
Progression-regression hypothesis, 419–421
Psychological refractoriness in tracking, 365, 389–391

- Pursuit
  - rotor, 361, 369, 392, 402–403
  - tracking, 361, 369–372
- Range effect in tracking, 365–366
- Reaction time, 497
- Recall
  - clustering in, 219
  - free, 218
  - procedure, 224
  - response, 218, 225, 238
  - R-S, 219–220
  - stimulus, 218, 225
  - *see also* Retention
- Receptor input control, 459–464, 473
- Recognition, multiple choice, 228, 230–231
- Reflex
  - cybernetic nature of, 2–3, 429–431, 487
    - *see also* History of skill
- Reinforcement, 255–259, 308–310, 505
  - and nonreinforcement, 79
  - compared with feedback, 47–48, 56, 63–73, 77–81, 85, 445–446, 466, 470–471
  - in selective learning, 47–48, 56, 63–73, 77–81, 85
  - intermittent, 472
  - *see also* Feedback
- Reminiscence, 9–10, 12–14, 19–24, 402, 507
- Response(s)
  - amplitude, 322, 335
  - definition, 484
  - hierarchy, 252
  - in selective learning, 47, 56–57, 61, 63–66, 70–72, 84–85
  - integration, 494
  - interference, 230
  - learning, 226, 298–299, 496, 499–500
  - meaningfulness, 222–226
  - measures, 48–49, 57, 279–280, 318–319, 323–325, 335, 357–358
  - multidimensionality of, 429, 434–436, 447, 458, 474–475
  - multiple measures, 246
  - -produced cues, 271
  - recall, *see* Recall
  - requirements of, 509
  - similarity, 220–226
- Retention, 124–127, 217–218, 220, 224, 232–233, 264, 267–268, 315–360
  - artifacts of measurement, 352–358
  - association
    - categories of patterns, 338–340
    - clustering, 336–337
    - conflicting, 510
    - control-display stereotypes, 346
    - hierarchy, 321, 337–340, 346, 510–511
    - interassociations, 345
    - mediation, 345–358
    - normative rates and networks, 336, 340–345
    - overlap, 338
    - pre-experimental, 340–345, 510–511
    - stimulated recall, 252, 333–347, 358–359
    - strength, structure, and rules, 251–253, 318–319, 333–347, 358–359
  - correlation
    - coefficient of determination, 328, 356
    - item analysis, 344–345
    - matrix or pattern analysis, 124–127, 158, 330–332, 344–345
    - measure of, 323–327, 335, 351–358
  - dichotomous score, 318–319, 335, 359
  - environments, 333–335
  - errors
    - blanks, importations, other list, 341
    - intrusions, 333–347
  - forgetting
    - accumulation of variance, 330–332, 344–345
    - control over, 335
      - *see also* Retention, reliability model
  - frequency hypothesis, 337, 511
  - individual differences, 163–164, 345, 354–358, 512
  - interpolated activity, 328–329, 345
  - learning
    - compared with repetition, 331–332
    - compared with retention, 321, 347, 353–357
    - complication or simplification process, 332
    - response and stimulus, 334
  - matrix analysis, *see* Retention, correlation
  - measures of, 317–325

  correlation, mean, and variance, 323–332, 353–358
  critical discussion, 351–358, 511–514
methods
  anticipation, 317–318
  controlled recall, 327, 333–347
  disengaging procedure, 328–329
  free association, 337–345
  free recall, 315–333, 337, 342–346
  overlearning, 317–318
  paired associates, 337
  recognition, 317–318, 336
  relearning, 317–321
  repeated recalling, 125–127, 330–332
  repetition, 316, 322–325, 331
    of response, response alternative, and feedback, 322–323, 333–335
  stimulated recall, 252, 333–347, 358–359
misrecall, 340–345, *see also* Retention, errors
motor tasks compared with verbal, 17–19, 333, 345–346, 509–511
performance distinction, 319
positioning a lever task, 320–323, 335, 355
proaction, 316–317, 329–330, 341
reliability (consistency) model, 327–330, 357–358
reorganization, 337
retroaction, 329–330
rote memory, 329
set, 329, 355, 357
short-term, 219, 329, 355
stimulated or reminded recall
    events, 252, 333–347, 358–359
  reinstatement of memory, 333, 340
structure, *see* Retention, correlation
task analysis, 321, 334–335
temporal course, 315–333, 342–345, 353–359
theory, *see* Retention, reliability model
threshold, 318, 358
trial
  alternative events and feedback, 322–323
  events or components of, 320–323, 333–335, 353
  relations between and within, 322, 346
  *see also* Recall
Reticular formation, 171–172
Reversibility, 215–216
Scales
  equal interval scales in retention, 353, 356, 358
  response versus feedback, 279–280
Selective learning, 47–107, 261, 268, 273–274, 490, 494, 508
  activity, 99–107
  apparatus, 49–55, 89
    Complex Coordinator, 50–51, 55, 74, 85
    Discrimination Reaction Timer, 50–52, 86–89
    Finger Maze, 50–51, 55, 57, 72, 87–89
    Multiple Serial Discrimeter, 50–52, 55, 76–78, 91, 103–106
    Selective Mathometer, 50, 52–55, 61, 67, 73–75, 77–89
    Star Discrimeter, 50, 52, 84–85
  behavior process, 103–106
  degrees of solubility, 81
  dynamic process, 100–103
  effect, law of, 48, 79
  equation (approximation formula), 59–60, 99–101
  extinction, 79–80
  facilitation, 83–86
    learning-to-learn, 83–84
    proactive, 84–85
    retroactive, 84
  habit reversal, 54, 83
  interference, retroactive, 83–86
  intervening variable
    habit, 72
  learning
    cue-reversal-discrimination, 72–73
    discrimination, 57, 72–73
    interpolated, 83–84
    mastery (overlearning) and retention, 72
    one trial, 72
    relearning trials, 85
    selective, 52–54, 62–63, 71, 76, 81, 85–86
      complexity of task, 73–74
      compound, 57

      descriptive analysis of, 54–56
      four basic elements, 55–56
      laws of, 73–89
      simple, 57
      spatial, 57–58
         generalization, 85–86
         position, 78
      theory of, 48–49, 56–73, 99–101
   set, 72, 83
   theories of, 57–58, 99–101
   trial-and-error, 51, 57, 63, 77, 79
   without awareness, 105–106
multiple-choice situations, 47–49, 89
nonreinforcement, 79
ongoing activity, 101–103
organismic variables, 60, 86–89, 101–106
   categories
      age-in-years, 60, 87–88
      asymptotic capacity, 60, 81
      habit aptitude, 60, 86–87
      individual drive, 60, 88–89
      perceptual-motor skill variability, 55, 81–82
      sex, 60, 86–88
   individual differences, 86–89
reinforcement, 47, 63–73
   serial, 77–78
   S-R theory, 56–57
      contiguity, 71–72, 85
   terminal, 77–78
   *see also* Feedback
repetition, 47
response, 47, 85
   generalization, 84
   postremity, 65–73
   probabilities, 55–56, 61
   selection, 57
   variables
      amplitude, 56
      frequency, 63–67, 70–72
stimulus, 85
   complexity of task, 60, 73–77
   reinforcing-feedback contingencies, 52, 60, 77–81
   selection, 57
   transfer-of-training, 83–86
   work distribution, 60, 81–83
stimulus-response, 101–106
   tendency, 102–103
   transfer of training, 60, 62, 76, 83–86
   trial, defined, 101, 104–105
   verbal control, 103–106
Short-term memory, 307–308, 327–330
Signal detection, 202–204, *see also* Attention, Monitoring, Vigilance, Watchkeeping
Similarity
   concept, 456–459
   conceptual, 498
   intertask, 220–226, 235
   intralist, 493
   intratask, 226–233
   paradox (in transfer), 221–222
   task, 507
Simplification (learning as a process of), 120–121, 136, 158–163
Skaggs-Robinson hypothesis, 220
Span
   of apprehension, 205
   of attention, 205
Spatial organization of behavior, 471
   and response specificity, 450, 454–456, 475
   axial reference mechanisms, 450–452, 483–484
   differential detection, 447, 450–454
   versus temporal, 446–447
Speech
   and feedback control, 438, 442, 468–469, 476–477
   and paired-associate learning, 500–501
Spontaneous recovery, 504, 508
Stimulus
   definition, 345–347, 484
   equivalence, 228
   functional, 225
   generalization, 85–86, 253
   in selective learning, 52, 57, 60, 73–86
   interference, 230
   learning, 226, 230
   meaningfulness, 222–226
   predifferentiation, *see* Differentiation
   selection, 225–226, 492
   similarity, 220–226
   variables in tracking, *see* Tracking, display
Stimulus-response (S-R)

- compatibility, 511
- psychology, 56–58, 63–73, 484, 487, 492, 498
- String model, 141–143
- Superdiagonal forms
  - in practice, 112–114, 158, 357
  - in recall, 124–127
  - in retention, 330–332
  - interpretation of, 120–123
  - rough, 123
- Symbolic
  - analogue, 302
  - learning, 468–469, 476–477
- Synthesis, *see* Composition rules, analysis and synthesis
- Target, 369–380
  - on target indicator, 278
  - size, 280–283
  - *see also* Feedback, transformations, Tracking, display
- Task(s)
  - classification, 131–133, 150–158
  - composition, 133–136
  - difficulty, 123
  - learning, *see* Apparatus and tasks
  - motor and verbal compared, 489–516
  - similarity, 131
  - *see also* Motor, tasks
- Teaching machines, 268
- Television techniques in cybernetics, 444–452, 465–467
- Tetrad differences, 114, 135
- Theory in verbal and motor learning, 504–509
- Tool using feedback, 464–469
- Trace theory, 263–264, 302
- Tracking, 261–262, 276–280, 287–288, 299, 361–424
  - acquisition, 361
  - anticipation in, 183–187
  - challenge effect in, 392–393, 403
  - compatibility of control with display, 381–382
  - compensatory, 361, 369–372, 381, 386, 394
  - computer model of, 179–180
  - contour, 361, 390–391
  - control, 380–384
    - "feel," 382–384
    - gain, 382, 386–387, 417–418
    - joystick, 380–382
    - pressure, 383
    - resistance, 382–384
    - rolling ball, 381
    - stylus, 380–381
  - control system, 384–389
    - acceleration, 383–384, 386–387, 392–393
    - aided, 384–385, 390, 413, 417
    - degradation, 388–389
    - differentiating, 381–384
    - intermittent, 388
    - "jerk," 384
    - maximum output, 387–388
    - order, 384–385
    - oscillation, 388–389
    - positional, 383–385, 387, 392–393
    - rate, 381, 383–387, 413
    - time lags, 369, 385–388
      - exponential, 385–386, 416–418
      - sigmoid, 386, 418
      - transmission, 385, 416–417
  - display, 369–380
    - aiding, 415
    - auditory, 378–379
      - pitch, 379
      - three dimensional, 379
    - augmented, 400–401
    - brightness, 377–378
    - degradation, 374–375, 400
    - delayed, 376
    - filtered, 376
    - flashing light, 377
    - intermittent, 374–375
    - loudness, 378–379
    - magnification, 371–372, 374, 393
    - multiple, 376–377
    - noisy, 375–376
    - predictor, 373, 387, 416
    - pursuit-compensatory, 411–414
    - quickened, 373, 381, 387, 414–416
    - sensory dimension, 377–380
    - vibratory, 378–379
    - visual, 377–378
      - depth, 378
      - length, 377
      - movement, 378
  - dither in, 396, 418–419

error peaks in, 179–181
eye movements and, 175, 178–179
man as tracker, 389–398
  describing function, 393–398
  first-order lag, 395
  first-order lead, 395
  indifference threshold, 395
  intermittency, 389–391, 396–398
  neuromuscular lag, 395
  non-linear transfer, 396
  non-steady behavior, 396–397
  reaction-time lag, 369, 389–392, 394–395, 397
  remnant of transfer function, 394, 396–397
  response gain, 394
  spare capacity, 392–393
  strategies, 390–392, 394, 397–398
  stress induced by task, 392–393
  transfer function, 393–398
  variability, 397
tracks, 364–369
  circular, 369
  multiple, 368–369
  ramp function, 365–366
  sinusoidal, 366–368
    multiple sine wave tracks, 368
    single sine wave tracks, 366–367
  step functions, 361, 364–365
    isolated steps, 364
    pairs of steps, 364–365
    series of steps, 365
training in tracking, 398–403
  distribution of practice, 402–403
  dynamic, 400
  progressive part, 398–399, 421–422
  reminiscence, 402
  simplification of task, 398–401
  task complexity, 421–422
  task organization, 421–422
  transfer, 372, 385, 398–400
    asymmetrical, 372, 459
  warm-up in, 402
two-dimensional, 368–369, 377
Training
  correlation patterns in, 141–143
  factor analysis in, 139–140
  pilot, 139–143
Transfer functions, 393–398
Transfer of training, 123–124, 215–254, 274–278, 288, 372, 385, 398–400, 454–459
  and learning-performance distinction, 251, 319
  asymmetrical, 372, 459
  bilateral, 507
  cybernetic interpretation, 456–459, 475
  difficulty in, 235–236
  displaced feedback and, 454–456
  in selective learning, 83–86
  intratask, 83, 217, 226–233
  measures, 248–249
  motor skills, 215, 233–238, 245–247
  paradigms (A-C, A-Br, C-B, C-D), 222–227
    in motor skills, 246–247
  surface (Osgood), 221
  task complexity in, 233
  verbal pretraining in, 233–244
  *see also* Facilitation, Interference, Whole-part
Trial(s)
  alternative events and feedback, 322–323
  compared with training, 139
  defined, 101, 104–105
  delay of information feedback, 268–269, 306–307
  discrete, 259
  distribution and retention, 321
  events in, 259, 262–263, 320–323, 333–337, 353
  relations between and within, 322, 346
Trial-and-error learning, *see* Selective learning
Twin studies, 129–131
Unlearning, 225, 332, 491
Variance, variability
  as forgetting index, 323–332, 356–358
  changes with practice, 115, 159–160, 282–283, 286–287
  in retention, 323–332, 345, 356–358, 511–514
  in tracking, 397
  specific, 143–144
Verbal
  learning, 304–308

- maze, 304–305
- mediation, 103–106
- satiation, 219
- tasks
  - compared with motor, 53–54, 233–239, 245–251, 298–299, 333, 345–346, 489–516

Vigilance, 170–171, 187–190, 202–204, 509
- and arousal theory, 209–210
- and attention, 208–209
- as a mechanism of response, 203
- tests, 203
- *see also* Attention, Monitoring, Signal detection, Watchkeeping

Warm-up, 18, 27–29, 510, 512

Watchkeeping, 202–204
- spatial uncertainty in, 206–208
- stimulus uncertainty in, 206–207
- task, 206–208
- temporal uncertainty in, 206–207
- *see also* Attention, Monitoring, Signal detection, Vigilance

Whole-part, 10–11, 25–27, 133–136, 139–143, 164, 398–399, 421–422, 497–498, 501–504, *see also* Analysis, component processes, Composition rules, analysis and synthesis, History of skill, Transfer of training

Wisconsin Card Sorting Test, 309

S.C.E.R.T., West Bengal
Library
Calcutta
25/3, B C. R

**Form No. 3.**

PSY, RES.L-1

**Bureau of Educational & Psychological Research Library.**

The book is to be returned within the date stamped last.

WBGP-59/60-5119C-5M

**Form No. 4**

**BOOK CARD**

Coll. No............. Accn. No......

Author.............................

Title..............................

| Date. | Issued to | Returned on |
|---|---|---|
| ............ | .......... | ............ |
| ............ | .......... | |